

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:15:08 ; Search time 163 Seconds

(without alignments)
1657.775 Million cell updates/sec

Title: US-10-661-430-1

Perfect score: 2031

Sequence: 1 MAFRLAVARLKLTLVLCNV.....VDLKKDCRRLRDPPTCF 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2031	100.0	383	2	O9GUM2_CAEEL	O9GUM2 caenorhabdi
2	1848.5	91.0	384	2	O60PPO_CABER	O60PPO trichoplusi
3	775.5	38.2	421	2	O6J4T9_TIRNI	O6J4T9 trichoplusi
4	743	36.6	260	2	O7Q2R7_ANOCA	O7Q2R7 anopheles g
5	736.5	35.3	498	2	O6NU80_XENIA	O6NU80 xenopus lae
6	730	35.9	498	2	O8AVL9_XENIA	O8AVL9 xenopus lae
7	729	35.9	336	2	O6DDG6_XENIA	O6DDG6 xenopus lae
8	729	35.9	354	2	O6DIT3_XENTR	O6DIT3 xenopus tro
9	725	35.7	395	1	B4GT3_MOUSE	B4GT3 xenopus tro
10	720	35.5	354	2	O5PPI6_XENIA	O5PPI6 xenopus lae
11	720	35.5	396	1	B4GT3_BOVIN	B4GT3 xenopus lae
12	717	35.3	393	1	B4GT3_HUMAN	B4GT3 xenopus lae
13	717	35.3	393	1	B4GT3_PONRY	B4GT3 xenopus lae
14	717	35.3	395	1	B4GT3_CRIGR	B4GT3 xenopus lae
15	716	35.3	395	1	B4GT3_RAT	B4GT3 xenopus lae
16	709	34.9	403	2	O9XZ05_DROME	O9XZ05 drosophila
17	696	34.3	360	2	O4T5R7_TENG	O4T5R7 tetraodon n
18	694	34.2	395	2	O71VA3_HUMAN	O71VA3 homo sapien
19	680.5	33.5	362	2	O92074_CHICK	O92074 gallus galli
20	679.5	33.5	341	2	O4SEF2_TENG	O4SEF2 tetraodon n
21	671.5	33.1	393	2	O99PCL_CRIGR	O99PCL cricetus
22	670	33.0	373	2	O92073_CHICK	O92073 gallus galli
23	667.5	32.9	399	1	B4GT1_MOUSE	B4GT1 xenopus lae
24	666.5	32.8	413	2	O9NDP2_CIOIN	O9NDP2 ciona intes
25	666	32.8	397	1	B4GT1_HUMAN	B4GT1 xenopus lae
26	660	32.5	372	2	O4V9L9_HUMAN	O4V9L9 homo sapien
27	657	32.3	372	1	B4GT2_HUMAN	B4GT2 xenopus lae
28	655.5	32.3	353	2	O5UJQ4_BRARE	O5UJQ4 brachydanio
29	651.5	32.1	402	1	B4GT1_BOVIN	B4GT1 xenopus lae
30	651	32.1	344	1	B4GT4_CRIGR	B4GT4 cricetus
31	648.5	31.9	369	1	B4GT2_CRIGR	B4GT2 cricetus

32	646.5	31.8	369	1	B4GT2_MOUSE	O92Y2 m beta-1,4-
33	642.5	31.6	344	1	B4GT4_HUMAN	O60513 h beta-1,4-
34	635	31.3	344	1	B4GT4_RAT	O661h1 r beta-1,4-
35	633	31.2	344	1	B4GT4_MOUSE	O9J104 m beta-1,4-
36	632	31.1	430	2	O4SY57_TENG	O4SY57 tetraodon n
37	632	31.1	557	2	O4Y16_TENG	O4Y16 tetraodon n
38	629	31.0	310	2	O64U7_PIG	O64U7 sus scrofa
39	621.5	30.6	347	2	O6GLP1_XENIA	O6GLP1 xenopus lae
40	616	30.3	319	2	O6BEO0_BRARE	O6BEO0 brachydanio
41	596.5	29.4	382	2	O4RLE6_TENG	O4RLE6 tetraodon n
42	593	29.2	382	2	O5F3E5_CHICK	O5F3E5 gallus galli
43	584.5	28.8	388	1	B4GT5_MOUSE	O9Jmk0 m beta-1,4-
44	584.5	28.8	388	1	O80W6_CRIGR	O80W6 cricetus
45	584	28.8	382	2	O4SDX9_TENG	O4SDX9 tetraodon n

ALIGNMENTS

RESULT 1
ID O9GUM2_CAEEL PRELIMINARY; PRT; 383 AA.
AC O9GUM2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Bt (Bacillus thuringiensis) toxin resistant protein 4
DE (UDFGAINAC:GICNAC(beta)-R (beta)1,4-N-acetylglucosaminyltransferase)
DE (BRE-4).
GN Name=bre-4; ORFNames=Y7J3E7A.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodietinae; Caenorhabditis.
OK NCBI_TaxID=6239;
RN NM_011911.1
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).
[2]
NM_011911.1
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22218093; PubMed=12167666; DOI=10.1074/jbc.M206112200;
RA Kauer Z.S., Van Die I., Cummings R.D.,
RT "Molecular Cloning and Enzymatic Characterization of a UDP-
GALNAc:GALNAc-beta-1,4-N-Acetylglucosaminyltransferase from
Caenorhabditis elegans.";
RL J. Biol. Chem. 277:34924-34932(2002).
[3]
NM_011911.1
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22970005; PubMed=12944392; DOI=10.1074/jbc.M308142200;
RA Griffiths J.S., Huffman D.L., Whitacre J.L., Barrows B.D.,
RA Marroquin L.D., Muller R., Brown J.R., Hennes T., Esko J.D.,
RA Arojan R.V.,
RT "Resistance to a bacterial toxin is mediated by removal of a conserved
glycosylation pathway required for toxin-host interactions.";
RL J. Biol. Chem. 278:45594-45602(2003).
[4]
NM_011911.1
RP NUCLEOTIDE SEQUENCE.
RX Griffiths J.S., Huffman D.L., Arojan R.V.,
RT Subtilized (UAN-2004) to the EMBL/Genbank/DBJ databases.
RL EMBL; AC005727; AAC33384.1; -; Genomic DNA.
DR EMBL; AY130767; AAM95168.1; -; mRNA.
DR HSSP; P08037; 100R.
DR Ensemble; Y7J3E7A.7; Caenorhabditis elegans.
DR WormBase; WBGenome0000269; bre-4.
DR WormPeP; Y7J3E7A.7; CE26412.
GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR003859; Galactosyl_T_2.

DR InterPro: IPR001998; Xylose_isom.
DR Pfam: PR02709; Galactosyl_T_2; 1.
DR PROSITE: PS00173; Xylose_isomerase_2; UNKNOWN_1.
KW Complete proteome; transferase.
SQ SEQUENCE 383 AA; 43914 MW; ADEDF44275B2CE81 CRC64;
Query Match 100.0%; Score 2031; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 3,4e-165;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFRLAVARLSKSLVLCVALLVHAMIKYKPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
DB 1 MAFRLAVARLSKSLVLCVALLVHAMIKYKPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
QY 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFQNPQPHLVGPIRVFLD 120
DB 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFQNPQPHLVGPIRVFLD 120
QY 121 PPKTLEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDRAHLRIMLHNLHSLAAKQQL 180
DB 121 PPKTLEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDRAHLRIMLHNLHSLAAKQQL 180
QY 181 VAFIVEQVANQTFNRGKLMNVGYDVASRLYPWCQCFIHDVDLLPEDDRNLVTCPIQPR 240
DB 181 VAFIVEQVANQTFNRGKLMNVGYDVASRLYPWCQCFIHDVDLLPEDDRNLVTCPIQPR 240
QY 241 MSVAIDKFNKYPYSAIFGGISALTQKHKLKINGFSDNDFMGWGEDDLATRTSMAGLV 300
DB 241 MSVAIDKFNKYPYSAIFGGISALTQKHKLKINGFSDNDFMGWGEDDLATRTSMAGLV 300
QY 301 SRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGQTKRMTRDGLSNLKYKLVNLEKPLY 360
DB 301 SRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGQTKRMTRDGLSNLKYKLVNLEKPLY 360
QY 361 RAVVDLLEKDCRRRLRDPPTCF 383
DB 361 RAVVDLLEKDCRRRLRDPPTCF 383
RESULT 2
Q60PPO CAEBR PRELIMINARY; PRT; 384 AA.
ID Q60PPO CAEBR PRELIMINARY; PRT; 384 AA.
AC Q60PPO;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DE Hypochemical protein CBG22165.
DE Hypochemical protein CBG22165.
GN Name=CBG22165;
OS Caenorhabditis briggsae.
OC Rhabdityota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; CAAC01000129; CAE74432.1; -; Genomic DNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR001998; Galactosyl_T_2.
DR InterPro: IPR001998; Xylose_isom.
DR Pfam; PF02709; Galactosyl_T_2; 1.
DR PROSITE; PS00173; Xylose_isomerase_2; UNKNOWN_1.
KW Hypochemical protein.
SQ SEQUENCE 384 AA; 43811 MW; 4F9BEC583A634B7C CRC64;
Query Match 91.0%; Score 1848.5; DB 2; Length 384;
Best Local Similarity 90.1%; Pred. No. 1.5e-149;
Matches 346; Conservative 16; Mismatches 21; Indels 1; Gaps 1;

QY 1 MAFRLAVARLSKSLVLCVALLVHAMIKYKPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
DB 1 MAFRLAVARLSKSLVLCVALLVHAMIKYKPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
QY 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFQNPQPHLVGPIRVFLD 119
DB 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFQNPQPHLVGPIRVFLD 120
QY 120 EPDFKLEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDRAHLRIMLHNLHSLAAKQQL 179
DB 120 EPDFKLEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDRAHLRIMLHNLHSLAAKQQL 180
QY 180 DVAIIVEQVANQTFNRGKLMNVGYDVASRLYPWCQCFIHDVDLLPEDDRNLVTCPIQPR 239
DB 180 DVAIIVEQVANQTFNRGKLMNVGYDVASRLYPWCQCFIHDVDLLPEDDRNLVTCPIQPR 240
QY 240 HNSVAIDKFNKYPYSAIFGGISALTQKHKLKINGFSDNDFMGWGEDDLATRTSMAGLV 299
DB 240 HNSVAIDKFNKYPYSAIFGGISALTQKHKLKINGFSDNDFMGWGEDDLATRTSMAGLV 300
QY 300 VSRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGQTKRMTRDGLSNLKYKLVNLEKPLY 359
DB 300 VSRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGQTKRMTRDGLSNLKYKLVNLEKPLY 360
QY 360 TRAVVDLLEKDCRRRLRDPPTCF 383
DB 360 TRAVVDLLEKDCRRRLRDPPTCF 384
RESULT 3
Q64T9 TRINI PRELIMINARY; PRT; 421 AA.
ID Q64T9 TRINI PRELIMINARY; PRT; 421 AA.
AC Q64T9;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE Beta 1,4-N-acetylgalactosaminyltransferase.
OS Trichoplusia ni (Cabbage looper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Dityraria; Noctuoidea;
OC Noctuidae; Plusiinae; Trichoplusia.
OX NCBI_TaxId=7111;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15173167; DOI=10.1074/jbc.M404925200;
RA Vadala N., Jarvis D.L.,
RT "Molecular Cloning and Functional Characterization of a Lepidopteran
Insect (beta)4-N-acetylgalactosaminyltransferase with Broad Substrate
Specificity, a Functional Role in Glycoprotein Biosynthesis, and a
RT Potential Functional Role in Glycolipid Biosynthesis.";
RL J. Biol. Chem. 279:33501-33518(2004).
DR EMBL; AY601103; AAT11926.1; -; mRNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR003859; Galactosyl_T_2.
DR InterPro: IPR001998; Xylose_isom.
DR Pfam; PR02709; Galactosyl_T_2; 1.
DR PROSITE; PS00173; Xylose_isomerase_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 421 AA; 48302 MW; EF3CC1168BBDAB CRC64;
Query Match 38.2%; Score 775.5; DB 2; Length 421;
Best Local Similarity 43.5%; Pred. No. 1.3e-57;
Matches 157; Conservative 63; Mismatches 116; Indels 25; Gaps 6;
QY 30 KIPSLY-----ENLTIGSSTLIADVDAMEAVLGNTASTSDLLDTWNSTFSPISSEV 80
DB 64 KIPSNYIEDSESYKINISLSNHTTASVYHPSSITETASLKDKNMTTODGAFAMISPT 123
QY 81 NQ--TFMEDIRPLIFPN-----QTLQFQNPQPHLVGPIRVFLDEPDKTLEKTY 130
DB 124 PLITIKLMSIKSYVTEDEGVKAAEVVTLPLCDSPMDL-GPIITNKTLELWEVKKF 182

```
QY 131 PBTAGHGMPKDCVARRHVAIIVPRDEAHLRLIMLHSLAKQQLDVAIFIVEQVA 190
DB 133 FEVEGGRSPNCPNCRARRHVAIIVPRDEAHLRLIMLHSLAKQQLDVAIFIVEQVA 242
QY 191 NOTFRGKLMNNGVYASRLYF--WQCFIFHDVLLIPEDRRLYTCPIOPRMSVAIDKF 248
DB 243 NMDFRKALMNGVYASRLYF--WQCFIFHDVLLIPEDRRLYTCPIOPRMSVAIDKF 302
QY 249 NYKLYPSAIFGSIATKDKHLKKGINSDFWGMGEGEDDDLLATRTSMGLKTSRYPTOLA 308
DB 303 HPKLYPEDI FGVSAANTLEQFTRVNGFSKTYMGWGGEDDDMSRLKKNYHIAKTKMSIA 362
QY 309 RYKMKHSTKATNPVNVKCRKYMGTQKRRRTDGLSNLKYKLVNLEKLYTRAVVDLE 368
DB 363 RYAMLDHKSTPNPK---RYGLLSQTSKTFQDGLSTLELYLVQVYHLIHLINIDE 419
QY 369 K 369
DB 420 R 420

RESULT 4
QY Q7Q2R7_ANOGA PRELIMINARY; PRT; 260 AA.
ID Q7Q2R7;
AC Q7Q2R7;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE ENSANGF0000010877 (fragment).
GN ORFNames=ENSANG0000008388;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anophelinae.
NCBI_Taxid=180454;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG The Anopheles gambiae re-annotation.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -i- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008968; EMBL3217.2; Genomic_DNA.
DR InterPro: IPR003959; Galactosyl_T_2.
DR InterPro: IPR001998; Xylose_Isom.
DR Pfam: PF02709; Galactosyl_T_2; 1.
DR PROSITE: PS00173; XYLOSE_ISOMERASE_2; UNKNOWN_1.
FT NON TER 1 1
FT NON TER 260 260
SQ SEQUENCE 260 AA; 29720 MW; 24B8F310B4781932 CRC64;

Query Match 36.6%; Score 743; DB 2; Length 260;
Best Local Similarity 54.3%; Pred. No. 4e-55;
Matches 139; Conservative 36; Mismatches 77; Indels 4; Gaps 2;
```

```
QY 291 TRTSMAGLKVSRYPQIARYKMKHSTKATNPVNVKCRKYMGTQKRRRTDGLSNLKYK 350
DB 181 NRKLVGHSHIARYPNINARIYMLSHKKEKANPK---RREKLVNAGKRRDSGLSLHQL 237
QY 351 VNLKRPVLYTRAVVDL 366
DB 238 VNLKRPVLYTRAVVDL 253

RESULT 5
QY Q6N80_XENLA PRELIMINARY; PRT; 498 AA.
ID Q6N80;
AC Q6N80;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE MGCB1163 protein.
GN Name=MGCB1163;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_Taxid=8355;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., USCIN T.B., Toshynski S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzanski M.I., Skalski U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RA initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC Klein S., Strausberg R.;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC068719; AAH68719.1; mRNA.
DR GO: GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro: IPR003859; Galactosyl_T_2.
DR Pfam: PF02709; Galactosyl_T_2; 1.
SQ SEQUENCE 498 AA; 56855 MW; 5A322F40FED0C01F CRC64;

Query Match 36.3%; Score 736.5; DB 2; Length 498;
Best Local Similarity 44.6%; Pred. No. 3.5e-54;
Matches 149; Conservative 54; Mismatches 114; Indels 17; Gaps 6;
```


RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnatratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Oocytes;
 RA Klein S., Gernard D.S.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC077601; AAH77601.1; -; mRNA.
 DR GO: GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro: IPR003859; Galactosyl_T_2.
 DR InterPro: IPR001998; Xylose_isom.
 DR Pfam: PF02709; Galactosyl_T_2; 1.
 DR PROSITE: PS00173; XYLOSE ISOMERASE 2; UNKNOWN 1.
 SQ SEQUENCE 336 AA; 38924 MW; 8E1A73D4CDD8F0C CRC64;

Query Match 35.9%; Score 729; DB 2; Length 336;
 Best Local Similarity 43.6%; Pred. No. 9.1e-54;
 Matches 137; Conservative 62; Mismatches 95; Indels 20; Gaps 4;

QY 61 STSDLDLDMN-----STFSPISFVNQSFMEDIRILFPDNOQLQFCNQTPHLYGPIR 115
 DB 25 SASMGILNHNSQAGDVANLSLSPSLR-----ELPYCEPTSPFJGGPIR 74
 QY 116 VFLDEPDEFKLEK---YPTDTHAGHGMPKDCVARRHVAIIVPYDRERHRLMNLHLS 172
 DB 75 VGL--PENLLLEVERKKNPYVYKGRKRPDCSESTHKTAVITPHNRREHLKLYLYLHP 132
 QY 173 LLAQQQLDYAIFIVGVAQNTFRGKLMNVGYDVASRLYPMOCFLPHDVLLPEDDRNLY 232
 DB 133 FLQROQLNGYIYIIHQAGNFTFRRAKLINVGFEAMKEDMDCLFPHVDVLLPEDDRNLY 192
 QY 233 TPCIPQPHMSVAIDENYKLPYSALFGISALTKHLLKKGESNDPMWGGEEDDLAR 292
 DB 193 ICDRPPKASISAMDKFGYLPYKSYFGVSALSPQYMKMGFPNNYMGWGGEEDDIDIGR 252
 QY 293 TSMAGLKVSRYPTQIARVYKMGKSTGATNPVVKCRKYGKIGQTKRRTDGLSNLKYLAVN 352
 DB 253 VALSGMLISRPVSQYGRYKMGKSHGDKNGEONPKRPMNLTKTRTKQDGMALQYILIS 312
 QY 353 LELKPLLYTRAVDL 366
 DB 313 KELQPLTYITIVDI 326

RESULT 8
 Q6DIT3_XENTR
 ID Q6DIT3_XENTR PRELIMINARY; PRT; 354 AA.
 AC Q6DIT3;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE UDP-Gal:betaGalCNAC beta 1,4-galactosyltransferase, polypeptide 3.
 GN Name=B4gal13-Prot;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus; Silurana.
 OK NCBI_TaxID=8364;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=whole body;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Haptschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnatratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=whole body;
 RA Klein S., Gernard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC075452; AAH75452.1; -; mRNA.
 DR GO: GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro: IPR003859; Galactosyl_T_2.
 DR Pfam: PF02709; Galactosyl_T_2; 1.
 DR GlycoSyL:transferase; Transferase.
 SQ SEQUENCE 354 AA; 40943 MW; 761A22CB3DDF072 CRC64;

Query Match 35.9%; Score 729; DB 2; Length 354;
 Best Local Similarity 39.7%; Pred. No. 9.7e-54;
 Matches 145; Conservative 64; Mismatches 122; Indels 34; Gaps 5;

QY 8 VARLSLVLCANLVLLVYAMTYK--IPSLYENLTIGSSLTLDADVAMENAVLNTASTSD 65
 DB 6 IDRRFLLLLFSSQLFLFLILYRRGTAGIFQGLFGSGD----- 43
 QY 66 LLDTNSTFSPISFVNQSFMEDIRILFPDNOQLQFCNQTPHLYGPIR-VFLDEPDK 124
 DB 44 --PFWDYRTADVYTNLSLFPD-----PENVNSQYCPAKSPILVGLSPSFFHNPPIK 95
 QY 125 TLEKIYPTDTHAGHGMPKDCVARRHVAIIVPYDRERHRLMNLHLSLAKQQLDYAIF 184
 DB 96 KYQQGNRYVYKGYSPRHCGRYRTAVIIPYRNEPHRLTYLYLHPFLQQLHYAVF 155
 QY 185 IYEVYANQTFNRGKLMNVGYDVASRLYPMOCFLPHDVLLPEDDRNLYTTCIQ-PRHNSV 243
 DB 156 IVHQAGNGTFRRAKLINVGFEALKDMDCLVLDVLPENDVNLVICOEEYKHLAS 215
 QY 244 AIDKENYKLPYSALFGISALTKHLLKKGESNDPMWGGEEDDLARTSMAGLKVSRY 303
 DB 216 AMDKFDYSLPYTYTYGVSALTTPHYKMGKSHGDKNGEONPKRPMNLTKTRTKQDGMALQYILIS 275
 QY 304 PTQIARVYKMGKSTGATNPVVKCRKYGKIGQTKRRTDGLSNLKYLAVNLEIKPLLYTRAV 363
 DB 276 PLSVGRYKMGKSHGDKNGEONPKRPMNLTKTRTKQDGMALQYILIS 335
 QY 364 VDILE 368
 DB 336 VDICE 340

RESULT 9
 B4GT3_MOUSE
 ID B4GT3_MOUSE STANDARD; PRT; 395 AA.
 AC Q91YY2; Q9QY13;
 DT 13-SEP-2005 (Rel. 48, Created)

DT 13-SBP-2005 (Rel. 48, Last sequence update)
 DT 13-SBP-2005 (Rel. 48, Last annotation update)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.1-) (Beta-1,4-GalTase 3)
 DE (Beta4Gal-13) (b4Gal-13) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-
 galactosyltransferase 3) [includes: N-acetylglucosamine synthase
 (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide
 beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-
 acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase
 (EC 2.4.1.1-)]
 DE Name=B4GalT3;
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RX MEDLINE=98259838; PubMed=9597550; DOI=10.1093/glycob/8.5.517;
 RA Lo N.-W., Shaper J.H., Peysner J., Shaper N.L.;
 RT "The expanding beta 4-galactosyltransferase gene family: messages from
 the databanks.";
 RL Glycobiology 8:517-526(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina G., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant P.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.S.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hulton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.E.N., Krzywinski M.I., Skalek U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Responsible for the synthesis of complex-type N-linked
 oligosaccharides in many glycoproteins as well as the carbohydrate
 moieties of glycolipids (By similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
 glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
 beta-D-glucosaminylglycopeptide.
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
 N-acetylglucosamine.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
 Golgi stack (By similarity).
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 7 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL: AF142671; AAF22221.1; -; mRNA.
 CC EMBL: BC013619; AAH13619.1; -; mRNA.
 DR HSSP: P08037; ITCX.
 DR Ensembl: ENSMUSG0000052423; Mus musculus.
 DR MGI: MGI:1928767; B4galT3.
 GO: GO:0016021; C:integral to membrane; TAS.

DR InterPro: IPR003859; Galactosyl T. 2.
 DR PANTHER: PTHR19300; Galactosyl T. 2; 1.
 DR Pfam: PF02709; Galactosyl T. 2; 1.
 KW Glycoprotein; Glycosyltransferase; Golgi stack; Manganese;
 KW Metal-binding; Multigene family; Signal-anchor; Transferase;
 KW Transmembrane.
 FT TOPO_DOM 1 10 Cytoplasmic (Potential).
 FT TRANSMEM 11 31 Signal-anchor for type II membrane
 FT FT protein (Potential).
 FT TOPO_DOM 32 395 Luminal (Potential).
 FT METAL 199 199 Manganese (By similarity).
 FT METAL 293 293 Manganese (By similarity).
 FT CARBOHYD 57 57 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 168 168 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 339 339 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 387 387 N-linked (GlcNAc...) (Potential).
 FT DISULFID 79 121 By similarity.
 FT DISULFID 192 211 By similarity.
 FT CONFLICT 353 353 S -> A (in Ref. 1).
 SQ SEQUENCE 395 AA; 44084 MW; 16401B033EB4B21 CRC64;
 Query Match 35.7%; Score 725; DB 1; Length 395;
 Best Local Similarity 48.9%; Pred. No. 2.5e-53;
 Matches 134; Conservative 44; Mismatches 94; Indels 2; Gaps 2;
 QY 95 PMNQTQPCNQTPPHLVGPIRV-FLDEPDKLTKIYPTTHAGHGMPDQVARRHVAII 153
 DB 71 PPAQALPYCPBESFPLVGVSFSFVPSLAELVEGNRAVESGGRPRAPGCEPSTAI 130
 QY 154 VPYRDEAHLRIHLNHLAKQQIDVAIFVEQVANOENRKGIMNTGYDVASFLYEW 213
 DB 133 VPRAREHHLRIILYHLHPLOQLAAGIYVHQNGTGNRAKLANTVGEALDEEWM 190
 QY 214 QCFIFPDVLLPDDENLQTC-PIQPRHMSVAILDKENYCLPSAIGGTSALTKDKLKKI 272
 DB 191 DCLFLHDVLLPDDNHLVPCDPRGHAVAVANKRGYSLPPQYGVSAITPDYLYKM 250
 QY 273 NGFSNDFMGGEEDDLARTSMAGLKVSRPTQIARKIKHSTKATPVNKCCKRYKNG 332
 DB 251 NGPFSNDFMGGEEDDLARTSMAGLKVSRPTQIARKIKHSTKATPVNKCCKRYKNG 330
 QY 333 QTRKRTDGLSLNLYKLVNLEIKPLYTRAVVDL 366
 DB 311 RTQNSWTQDGMNSLTFYRLARELGPLYTNTADI 344
 RESULT 10
 Q5PPY8_XENULA PRELIMINARY; PRT; 354 AA.
 AC Q5PPY8;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DE LOC436041 protein.
 GN Name=LOC436041;
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pypidae;
 OC Xenopodidae; Xenopus; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.W., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
 RA Diachenko L., Matrusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussidi T.B., Toshlyuk S., Carninci P., Plange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Millhys S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 (3)
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RA Klein S., Gerhard D.S.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC087436; A087436.1; -; mRNA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
 DR CO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR003859; Galactosyl_T_2.
 DR Pfam; PF02709; Galactosyl_T_2; 1.
 SQ SEQUENCE 354 AA; 40970 MW; 016170773E168208 CRC64;

Query Match 35.5%; Score 720; DB 2; Length 354;

Best Local Similarity 40.3%; Pred. No. 5.7e-53;

Matches 143; Conservative 61; Mismatches 117; Indels 34; Gaps 5;

14 LTVLCVLLVHMYK--IPSLYENLTIGSSSTLADVDVAMEAVLGNTASTSDLLDVTN 71
 12 LLLLCESQLLFLILYRRAAGVFOGLRGGSD-----PPMD 47
 72 STFSPISEVNGTSFMEDIRPILEPDNQLQFCNQPPHLYGPIRV-FLDEPDKLEKTY 130
 48 YEKTDHVVNLSVFN-----PENVNQYCPAKPIIVGPIVSFQPPPTLKRVQQRN 101
 131 PPTHAGCHMPKDCVARRHVAIIPYRDEAHKIMLNLHLSLAKQODVAIFIVEQVA 190
 102 RYKPCGGYSPRRCFGRYRTAVIIPRNRBAHLRTLLYLHPLOQQLAYALFTVHQAG 161
 191 NOTFNKGLMNGYDVASRLYPWQCFIPHDVLLPEDDRLNLYTCPIQ-PRHMSVAIDKEN 249
 162 NQTFNRKAKLNLGVREALKLDEMDCILHLDVLPENDNLYICDEBYRKLHLSAMDKEH 221
 250 YKLPSAIFGIGISALTQDHLKKINGESNDFWGWGEGEDDLDATRTSMAGLKVSRYPQTQAR 309
 222 YSLPYTVYGGVSALTPDQYMRINGEPNPGWGGEDDDIAMRIRLAGMSITRTPLSLGR 281
 310 YKMGKSTATNPVAKCRKRYKMGQTKRKRTRDGLSLUKLVNLEKPIYTAADV 364
 282 YKMSIHNRDSGNENSKRYDOLGNTRTWRDEGMNSLDFKLISRTAPLTYITV 336

RESULT 11

BAGT3_BOVIN STANDARD; PRT; 396 AA.

AC Q5EA87; Q5E9K4;
 DT 13-SBP-2005 (Rel. 48, Created)
 DT 13-SBP-2005 (Rel. 48, Last sequence update)
 DT 13-SBP-2005 (Rel. 48, Last annotation update)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.1.-) (Beta-1,4-GalTase 3)
 DE (Beta4Gal-1T3) (b4Gal-T3) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,4-galactosyltransferase 3) (UDP-gal:beta-GlcNAc beta-1,4-
 galactosyltransferase 3) [Includes: N-acetylglucosamine synthase
 (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide

DE beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-
 DE acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase
 DE (EC 2.4.1.-).
 GN Name=BAGALT3;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RA Hathay G.P., Sonstegard T.S., Clawson M.L., Heaton M.P., Keele J.W.,
 RA Snelling W.M., Weidman R.T., Smith T.P.L.;
 RT "Sequencing and analysis of Bos taurus full-length insert cDNA
 clones.";
 RT Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Responsible for the synthesis of complex-type N-linked
 CC oligosaccharides in many glycoproteins as well as the carbohydrate
 CC moieties of glycolipids (By similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
 CC glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
 CC beta-D-glucosaminylglycopeptide.
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
 CC N-acetylglucosamine.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
 CC Golgi stack (By similarity).
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 7 family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; BT020682; AAX08699.1; -; mRNA.
 CC EMBL; BT020686; AAX08885.1; -; mRNA.
 CC EMBL; BT020916; AAX08933.1; -; mRNA.
 DR InterPro; IPR003859; Galactosyl_T_2.
 DR PANTHER; PTHR19300; Galactosyl_T_2; 1.
 DR Pfam; PF02709; Galactosyl_T_2; 1.
 KW Glycoprotein; Glycosyltransferase; Golgi stack; Manganese;
 KW Metal-binding; Multigene family; Signal-anchor; Transferrase;
 KW Transmembrane.
 FT FT 10
 FT TOPO_DOM 11 31
 FT TRANSMEM 1 1
 FT FT
 FT TOPO_DOM 32 396
 FT METAL 200 200
 FT METAL 294 294
 FT CARBOHYD 57 57
 FT CARBOHYD 169 169
 FT CARBOHYD 340 340
 FT CARBOHYD 388 388
 FT DISULFID 80 122
 FT DISULFID 193 212
 FT FT 82
 FT CONFLICT 82
 FT SEQUENCE 396 AA; 44371 MW; 8C2611ED7C26250D CRC64;

Query Match 35.5%; Score 720; DB 1; Length 396;

Best Local Similarity 48.6%; Pred. No. 6.7e-53;

Matches 136; Conservative 44; Mismatches 98; Indels 2; Gaps 2;

89 IRPILEPDNQLQFCNQPPHLYGPIRV-FLDEPDKLEKTYPTTHAGCHMPKDCVAR 147
 66 VAPGGLPAPGCLPYCPKSPILLVGPISVSFSPVSLAIVERNPRVPGGRPARCEBR 125
 148 HRVAIIPYRDEAHKIMLNLHLSLAKQODVAIFIVEQVANTFNKGLMNGYDVA 207
 126 STATAIVHRAEHHRLTLHLHPLOQQLAYGIYVHQNGTFRNRAKLNLGVREA 185
 208 SRLYPWQCFIPHDVLLPEDDRLNLYTC-PIOPRHSVAIDKENYKLPYSAIFGIGISALT 266

Db 186 LRDEBDC.FLHVDLLEPBNHNLVYCDPGRPHVAVAMNKFGXSLPYQFQGVSAALP 245
 Oy 267 DHAKTINGSNBDMWNGCGERDDLATRTSMAGLVSRYPQIARVKKIKISTEATNVNKC 326
 Db 246 DOYLKKNQFPNEVWGWGEGEDDLATRVRLAKMKISRPPIVSGHYKWKVKGXGNEENPH 305
 Oy 327 RYKMGOTKRRWRPDRGLSNLKYLVNLEKPLVTRAVVDL 366
 Db 306 RPDVLVTRTONSWTQDDMNLSLTQLLSRELGLPLTNITADI 345
 RESULT 12
 B4GRT3 HUMAN STANDARD: PRT; 393 AA.
 ID B4GRT3 HUMAN STANDARD: PRT; 393 AA.
 AC 060512; 060910; Q9BPR2; Q9H8T2;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 48, Last annotation update)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-GalTase 3)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-GalTase 3)
 DE (Beta4Gal-T3) (b4Gal-T3) (UDP-galactose:beta-N-acetylglucosamine beta-1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-galactosyltransferase 3) (Includes: N-acetylglucosamine synthase
 DE (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide
 DE beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase
 DE (EC 2.4.1.-).
 GN Name=B4GRT3;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RX MEDLINE=98070356; PubMed=9405390; DOI=10.1074/jbc.272.51.31979;
 RA Almeida R., Amado M., David L., Levery S.B., Holmes E.H., Merx G.,
 RA van Kessel A.G., Rygaard E., Hassan H., Bennett E., Clausen H.;
 RT "A family of human beta4-galactosyltransferases. Cloning and
 RT expression of two novel UDP-galactose:beta-n-acetylglucosamine beta1,
 RT 4-galactosyltransferases, beta4Gal-72 and beta4Gal-75.";
 RL J. Biol. Chem. 272:31979-31991(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RX MEDLINE=98259838; PubMed=9597550; DOI=10.1093/glycob/8.5.517;
 RA Lo N.-W., Shaper J.H., Peyssner J., Shaper N.L.;
 RT "The expanding beta 4-galactosyltransferase gene family: messages from
 RT the databanks.";
 RL Glycobiology 8:517-526(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND BIOPHYSICOCHEMICAL
 RP PROPERTIES.
 RX MEDLINE=21472272; PubMed=11588157; DOI=10.1093/glycob/11.10.813;
 RA Guo S., Sato T., Shirane K., Furukawa K.;
 RT "Galactosylation of N-linked oligosaccharides by human beta-1,4-
 RT galactosyltransferases I, II, III, IV, V, and VI expressed in Sf-9
 RT cells.";
 RL Glycobiology 11:813-820(2001).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RG Human chromosome 1 international sequencing consortium;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
 RC TISSUE=Ovary;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,

RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibaishi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimeta M., Watanabe M., Hirooka S., Chiba Y.,
 RA Ieshida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Houta T.,
 RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togliya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuki H., Oshima A., Saeki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitani Y.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtsuki R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Ohtsuna K., Nakajima Y., Mizuno T., Moriyaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [6]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RN TISSUE=Brain, Eye, and Lymph;
 RC TISSUE=Brain, Eye, and Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Utsid T.B., Toshynki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Malhotra S.J.,
 RA Bosak S.A., McManus P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Holton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP REVIEW.
 RX MEDLINE=20047730; PubMed=10580128; DOI=10.1016/S0304-4165(99)00168-3;
 RA Amado M., Almeida R., Schwientek T., Clausen H.;
 RT "Identification and characterization of large galactosyltransferase
 RT gene families: galactosyltransferases for all functions.";
 RL Biochim. Biophys. Acta 1473:35-53(1999).
 CC -1- FUNCTION: Responsible for the synthesis of complex-type N-linked
 CC oligosaccharides in many glycoproteins as well as the carbohydrate
 CC moieties of glycolipids.
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
 CC glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
 CC beta-D-glucosaminylglycopeptide.
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
 CC N-acetylglucosamine.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
 CC Kinetic parameters:
 CC Km=63 uM for GlcNAc-B-S-pNP;
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
 CC Golgi stack.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;

CC		Isoid=060512-1; Sequence=Displayed;
CC	Name=2;	
CC	Isoid=060512-2; Sequence=VSP_014106, VSP_014107;	
CC	Note=No experimental confirmation available;	
CC	-1- TISSUE SPECIFICITY: Found in various tissues. Highest expression	
CC	in placenta, prostate, testis, ovary, intestine and muscle, and in	
CC	fetal brain.	
CC	-1- SIMILARITY: Belongs to the glycosyltransferase 7 family.	
CC		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use as long as its content is in no way modified and this statement is not	
CC	removed.	
CC		
DR	EMBL; Y12509; CAAT7311.1; -, mRNA.	
DR	EMBL; AF038661; AAC39734.1; -, mRNA.	
DR	EMBL; AB024435; BAA75820.1; -, mRNA.	
DR	EMBL; AL590714; CAH72145.1; -, Genomic_DNA.	
DR	EMBL; AK023311; BAB14520.1; -, mRNA.	
DR	EMBL; BC000276; AAH00276.1; -, mRNA.	
DR	EMBL; BC006099; AAH06099.1; -, mRNA.	
DR	HSSP; P08037; INMM.	
DR	Ensembl; ENSG00000158650; Homo sapiens.	
DR	HGNC; HGNC:926; B4GALT3.	
DR	MTM; 604014; -.	
DR	GO; GO:0008378; F:galactosyltransferase activity; TAS.	
DR	InterPro; IPR003859; Galactosyl_T_2.	
DR	PANTHER; PTHR19300; Galactosyl_T_2; 1.	
DR	Pfam; Pf02709; Galactosyl_T_2_1.	
KW	Alternative splicing; Glycoprotein; Glycosyltransferase; Golgi stack;	
KW	Manganese; Metal-binding; Multigene family; Signal-anchor;	
KW	Transferase; Transmembrane.	
FT	TOPO_DOM 1 10	Cytoplasmic (potential).
FT	TRANSMEM 11 31	Signal-anchor for type II membrane protein (potential).
FT	TOPO_DOM 32 393	Luminal (potential).
FT	METAL 197 197	Manganese (By similarity).
FT	METAL 291 291	Manganese (By similarity).
FT	CARBOHYD 57 57	N-linked (GLcNAc . .) (Potential).
FT	CARBOHYD 166 166	N-linked (GLcNAc . .) (Potential).
FT	CARBOHYD 337 337	N-linked (GLcNAc . .) (Potential).
FT	CARBOHYD 385 385	N-linked (GLcNAc . .) (Potential).
FT	DISULFID 77 119	By similarity.
FT	DISULFID 190 209	By similarity.
FT	VASAPPLIC 116 145	PAGEPPSRRTAIIYDHAREHTLLLYHL -> PAALPPRA PLTAAPPAACMLNLTCHPEQMKKNI (in isoform 2).
FT		/FTID=VSP_014106.
FT	VASAPPLIC 146 393	Missing (in isoform 2).
FT		/FTID=VSP_014107.
FT	CONFLICT 64 64	A -> R (in Ref. 2).
FT	CONFLICT 112 112	G -> A (in Ref. 2).
FT	CONFLICT 223 223	K -> S (in Ref. 2).
SO	SEQUENCE 393 AA; 43928 MW; CA0BF955F955BD1F4 CRC64;	
Query Match	35.3%; Score 717; DB 1; Length 393;	
Best Local Similarity	48.9%; Pred. No. 1,2e-52;	
Matches 134; Conservative 43; Mismatches 95; Indels 2; Gaps 2;		
QY	95 PDNQTFQCNPTRPHLVGPRIYV-FLDEBDFETLKIEIYPDTTHAGGHMPKDCVAHRRAVII	153
Dd	69 PAPQGLGYCEPERSPFLVGPSVSFPVSLAEIYERNRVRVEPGRGYRPACGEPSRSRTAI	128
QY	154 VPVYEDREAHIRIMAHNHSLAKKOQLDYAFIFIVOVANQFNRRKLNNYGDVASRLYPW	213
Dd	129 VPHAREHHNLKLLYLHPPLIQROOLAVGIYVIHQAGGTFRKKLLNVGREGRLRDREW	188
QY	214 QCFITFDVDLLPBDRLNYTC-PLOPRHMSVALIKFYVKLPYSALFGISALTYKHDKKI	272
Dd	189 DCLFLHDVDILLPENDHNLYVEDGRPRHVANAANKFGVSLPYPPGVGSVALTLDQYLKM	248
QY	273 NGESNDIEWMGCGEBDDLATRTSMAGLKVSRPTIOIARYKMIKHSTEATPNVKNCRYYIMG	332

```

Db      249   NGEFNEHWNGSGEDDDIATRAVLAKMKISRPPTSVGHKWKVKHGDCKGNENPHRFLLV 308
Qy      333 QTKRWTRDGLSNLKYKLVMLELKPLYTRAVVDL 366
Db      309 RTQSMWTQDGWNNSLTYOQLAREIGPLVTNTRADI 342

RESULT 13
BAGT3 PONPY
ID      BAGT3 PONPY          STANDARD;           PRT;         393 AA.
AC      OSNVNT3;
DT      13-SEP-2005 (Rel. 48, Created)
DT      13-SEP-2005 (Rel. 48, Last sequence update)
DT      13-SEP-2005 (Rel. 48, Last annotation update)
DE      Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-Galnase 3)
DE      (BetaGal-T3) [d4Gal-T3] (UDP-galactose:beta-N-acetylglucosamine beta-
DE      1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-
DE      galactosyltransferase 3) (Includes: N-acetylglucosamine synthase
DE      (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide
DE      beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-
DE      acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase
DE      (EC 2.4.1.-)).
DN      Name=B4GALT3;
OS      Pongo pygmaeus (Orangutan).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC      Pongo.
OX      NCBI_TaxID=9660;
[ ]
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC      TISSUE=Brain cortex;
RG      The German CDNA consortium;
RL      Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: Responsible for the synthesis of complex-type N-linked
CC      oligosaccharides in many glycoproteins as well as the carbohydrate
CC      moieties of glycolipids (by similarity).
CC      -!- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
CC      glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
CC      beta-D-glucosaminylglycopeptide.
CC      -!- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
CC      N-acetylglucosamine. (By similarity).
CC      -!- COFACTOR: Manganese. (By similarity).
CC      -!- PATHWAY: Glycosylation.
CC      -!- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
CC      Golgi stack (By similarity).
CC      -!- SIMILARITY: Belongs to the glycosyltransferase 7 family.
-----
Cc      This Swiss-Prot entry is copyright. It is produced through a collaboration
Cc      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
Cc      the European Bioinformatics Institute. There are no restrictions on its
Cc      use as long as its content is in no way modified and this statement is not
Cc      removed.
-----
Cc      EMBL; CR925982; CA129630.1; -. mRNA.
Cc      InterPro; IPR003859; Galactosyl_T_2.
Cc      Pfam; PTRH19300; Galactosyl_T_2; 1.
DR      DR PFAM: PF02709; Galactosyl_T_2; 1.
KW      Glycoprotein; Glycosyltransferase; Golgi stack; Manganese;
KW      Metal-binding; Multigene family; Signal-anchor; Transferase;
KW      Transmembrane.
FT      TOPO_DOM        1       10    Cytoplasmic (Potential).
FT      TRANSMEM        11      31    Signal-anchor for type II membrane
FT                                     protein (Potential).
FT      METAL           197      197    Manganese (By similarity).
FT      METAL           291      291    Manganese (By similarity).
FT      CARBOHYD         57        57    N-linked (GlcNAc...) (Potential).
FT      CARBOHYD        166        166    N-linked (GlcNAc...) (Potential).
FT      CARBOHYD        337        337    N-linked (GlcNAc...) (Potential).
FT      CARBOHYD        385        385    N-linked (GlcNAc...) (Potential).
FT      DISULFID         77        119 By similarity.
FT      DISULFID        190        209 By similarity.

```

Seq	SEQUENCE	393 AA;	43922 MW;	CA0BP9562D32D14	CRC64;
Qy	Query Match	35.3%;	Score 717;	DB 1;	Length 393;
Bst	Best Local Similarity	48.9%;	Pred. No. 1,2e-52;		
Matches	134;	Conservative 43;	Mismatches 95;	Indels 2;	Gaps 2;
Qy	95 PDNQTLOFCNCTPPLHVGPIRV-FLDEPDFETLEKIYPTDHAGHGMPKXCVARHRAVIL	153			
Db	69 PAPGLPCPCPRSPPLVGVSPVSFSPVSLAEIYERNDRVPEPGGRYPACCEPSPRAIIL	128			
Qy	154 VPPYDREAHLEIMLNHLNLSLAKOOLDAPIEIVQVANAQTFNRGKLMNVGVDSVRLYPM	213			
Db	129 VPHARAREHLLDLVHLHPFLQROQLAVGIYVTHQAGNGFTNRKALNVGYRELRBEEH	188			
Qy	214 QCFIFHDVLLPEDDRMLYTC-PIQPRHMSVALDKENYKLYPSAIFGGISALTKHLIKTI	272			
Db	189 DCLFLHDVLLDPENDHNLVYCDPGRPHRAVAANMKFGSLPYPGYFGVSALTPDYQLKM	248			
Qy	273 NGFSNDFPMGCGEDDLDLTRISMAGLKYSRFTPTQIARYKMIKSTTEATNPNNKRYKMG	332			
Db	249 NGFPIEYMGWGEEDDIALTRVRLAGMKISRPPTSGYKQWVKGHDGKNEENPHRFDLLV	308			
Qy	333 QTKRRWRPDGSLNLYKYLWNLKPLRYRAVVDL	366			
Db	309 RTQNSWIDQGNMSLTYYQLARELGLPYNTIADI	342			
RESULT 14					
BAGT3	CRIGR				
ID	BAGT3_CRIGR	STANDARD;	PRT;	395 AA.	
AC	Q80WN8;				
DT	13-SEP-2005 (Rel. 48, Created)				
DT	13-SEP-2005 (Rel. 48, Last sequence update)				
DT	13-SEP-2005 (Rel. 48, Last annotation update)				
DE	Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-Galrase 3)				
DE	Beta-1,4-galactosyltransferase 3 (UDP-Gal:beta-D-glucNAc beta-1,4-				
DE	1,4-galactosyltransferase 3) (UDP-Gal:beta-D-glucNAc beta-1,4-				
DE	galactosyltransferase 3) [Includes: N-acetylglucosamine synthase				
DE	(EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide				
DE	beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-				
DE	acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase				
DE	(EC 2.4.1.-).)				
GN	Name=B4GALT3;				
OS	Citricellus griseus (Chinese hamster).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;				
OC	Muroidea; Cricetidae; Cricetinae; Citricellus.				
OX	NCBI_TaxID=10029;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [MRNA].				
RC	TISSUE=Ovary;				
RX	PubMed=14567896; DOI=10.1021/bi0353068;				
RA	Lee J., Park S.-H., Sundaram S., Raju T.S., Shaper N.L., Stanley P.;				
RT	"A mutation causing a reduced level of expression of six beta4-				
RT	galactosyltransferase genes is the basis of the Lec19 CHO				
RL	glycosylation mutant.";				
RL	Biochemistry 42:12349-12357(2003).				
CC	- FUNCTION: Responsible for the synthesis of complex-type N-linked				
CC	oligosaccharides in many glycoproteins as well as the carbohydrate				
CC	molecules of glycolipids (By similarity).				
CC	- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-				
CC	glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-				
CC	beta-D-glucosaminylglycopeptide.				
CC	- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +				
CC	N-acetylglucosamine.				
CC	- COFACTOR: Manganese (By similarity).				
CC	- PATHWAY: Glycosylation.				
CC	- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of				
CC	Golgi stack (By similarity).				
CC	- SIMILARITY: Belongs to the glycosyltransferase 7 family.				
CC	-----				
CC	This Swiss-Prot entry is copyrighted. It is produced through a collaboration -				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				

CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----

DR EMBL; AY115337; AAMW7196.1; -, mRNA.

DR HSSP; P08037; IFGX.

DR InterPro; IPP003859; Galactosyl_T_2.

DR PANTHER; PTHR19300; Galactosyl_T_2; 1.

DR Pfam; PF02709; Glycosyltransferase; Golgi stack; Manganese;
KW Glycoprotein; glycosyltransferase; 3

KW Metal-binding; Multigene family; Signal-anchor; Transferase;

KW Transmembrane.

FT FT TOPO_DOM 1 10 Cytoplasmic (Potential).

FT TRANSMEM 11 31 Signal-anchor for type II membrane protein (Potential).

FT FT TOPO_DOM 32 395 Lumenal (Potential).

FT METAL 199 199 Manganese (By similarity).

FT METAL 293 293 Manganese (By similarity).

FT CARBOHYD 57 57 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 168 168 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 339 339 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 387 387 N-linked (GlcNAc...) (Potential).

FT DISULFID 79 121 By similarity.

FT DISULFID 192 211 By similarity.

SQ SEQUENCE 395 AA; 44060 MW; 7FB9760DC6329680 CRC64;

Query Match 35.3%; Score 717; DB 1; Length 395;
Best Local Similarity 49.1%; Pred.No.1.2e-52;
Matches 133; Conservative 43; Mismatches 93; Indels 2; Gaps 2;

OY 98 QTLQPCNTPPHLYGPPIRV-FLDEPDPELTLEKIPDTTHAGHGMPKCVARHRAVITVPY 156
Db :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 74 EALPFCEPSRPPLTGVPSVSFPVSLAEIERNPRVEPGGRYPAGGEASRTAIIIVPH 133
Db :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
OY 157 RDREAHRLIMLAHLASHLAKQOLUYAFILFEVQNAVOTFNRRGLANMGVDVASRLPYMQCF 216
Db 134 RAREHHRLLLYLHPFLQRQQLAVGIYVHQAGNGMFNRALKLVNGVRELRLOBEWDCL 193
OY 217 IFHDVILLPEDRNLRYTC-PIQPRHMSVALDKENFKYLPYSAILFGSIALTYKHLLKTINGF 275
Db 194 FLHDVILLPEHDNLRYCDDPRGPRHVAANMKFGSYLPYPQFGVSALTDPDYLLKNMGF 253
OY 276 SNDPRKGCGEDDLATRSMAGLKYSRPPQIAIKYMIKSTTEATNPNNKKRYIMGTQ 335
Db 254 PNEYWGMEGEDDIATRVRLAGMKSIRPPTSVMGYKMVKHGDGKNEENPHRFDLLVRTQ 313
OY 336 RRWRDGLSNLKYYKLVLNELAPLTVRAVVDL 366
Db 314 NSWTODGNSSLTYQLAKELGPLTYTNITADI 344

RESULT 15

BAGTS_RAT STANDARD; PRT; 395 AA.

ID BAGTS_RAT STANDARD; PRT; 395 AA.

AC Q6P768;

DT 13-SEP-2005 (Rel. 48, Created)

DT 13-SEP-2005 (Rel. 48, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-Galtrae 3)

DE (BetaGal-T3) (UDP-galactose:beta-N-acetylglucosamine beta-1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-galactosyltransferase 3) [includes: B-N-acetylgalactosaminyl synthase (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylgalactosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-acetylgalactosaminyl-glycolipid beta-1,4-galactosyltransferase (EC 2.4.1.-)].

DE Name=B4galt3;

OS Rattus norvegicus (Rat).

OC Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Sakurajida; Euteeria; Euarchoctoglossa; Gilres; Rodentia; Sciurognathi;
OC Murioidea; Muridae; Murinae; Rattus.

RN NCBI_nxid=10116;

RN [1]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 18:38:45 ; Search time 6017 Seconds
(without alignments)
10883.107 Million cell updates/sec

Title: US-10-661-430-2
Perfect score: 1152
Sequence: 1 atgcttcctgcattcgtgc.....acttcacacgtgtttttag 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank1:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sce:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1152	100.0	1152	2	AY130767	AY130767 Caenorhab
2	1152	100.0	1152	2	AY533306	AY533306 Caenorhab
3	262	22.7	63747	2	AC025727	AC025727 Caenorhab
4	192.6	16.7	1839	6	CQ590796	CQ590796 Sequence
5	192.6	16.7	2269	2	AF132158	AF132158 Drosophila
6	192	16.7	1269	2	AY601103	AY601103 Trichopl
7	191	16.6	1189	6	AR506512	AR506512 Sequence
8	191	16.6	1879	2	AY095531	AY095531 Drosophila
9	186.8	16.2	2796	2	AB036858	AB036858 Clona int
10	181.4	15.7	2082	2	AK116794	AK116794 Clona int
11	171.8	14.9	2265	5	BC077601	BC077601 Xenopus 1
12	170.2	14.8	2681	5	BC068719	BC068719 Xenopus 1
13	169.2	14.7	1338	9	AY117537	AY117537 Cricetulu
14	168.4	14.6	2007	4	BC061812	BC061812 Rattus no
15	168.2	14.6	1921	4	BT020916	BT020916 Bos tauru
16	168.2	14.6	1927	4	BT020682	BT020682 Bos tauru
17	168.2	14.6	1949	4	BT020868	BT020868 Bos tauru
18	167.6	14.5	1182	8	HSUDP GAL	Y12509 Homo sapien

19	167.6	14.5	1193	8	AB024435	AB024435 Homo sapi
20	167.6	14.5	1912	8	BC009985	BC009985 Homo sapi
21	167.6	14.5	1934	8	BC006099	BC006099 Homo sapi
22	167.6	14.5	1938	6	BD206381	BD206381 Human nuc
23	167.6	14.5	1938	6	AR400696	AR400696 Sequence
24	167.6	14.5	1938	6	AX013200	AX013200 Sequence
25	166.4	14.4	4819	5	AJ851705	AJ851705 Gallus ga
26	166	14.4	1885	8	AK092493	AK092493 Homo sapi
27	166	14.4	1915	8	BC000276	BC000276 Homo sapi
28	166	14.4	1920	8	AP038661	AP038661 Homo sapi
29	166	14.4	2340	5	BC075452	BC075452 Xenopus t
30	165.8	14.4	1954	5	CR760293	CR760293 Xenopus t
31	164.4	14.3	2479	2	AK112685	AK112685 Clona int
32	163.2	14.2	2626	5	BC041742	BC041742 Xenopus 1
33	161.2	14.0	2147	8	AF069054	AF069054 Homo sapi
34	161.2	14.0	2289	8	AP038664	AP038664 Homo sapi
35	161.2	14.0	2663	5	CR926530	CR926530 Xenopus t
36	159.6	13.9	1177	8	AB024742	AB024742 Homo sapi
37	159.6	13.9	1198	8	BC074835	BC074835 Homo sapi
38	159.6	13.9	1198	8	BC074835	BC074835 Homo sapi
39	159.6	13.9	1329	5	BC087436	BC087436 Xenopus 1
40	159.6	13.9	1355	8	BC069642	BC069642 Homo sapi
41	159.6	13.9	3931	8	AF097159	AF097159 Homo sapi
42	158	13.7	1962	9	BC013619	BC013619 Mus muscu
43	157.4	13.7	2098	5	CR942597	CR942597 Xenopus t
44	157.2	13.6	1853	6	BD191505	BD191505 Secrete
45	157	13.6	1714	8	BC004523	BC004523 Homo sapi

ALIGNMENTS

RESULT 1	AY130767	1152 bp	mRNA	linear	INV 17-SEP-2002
LOCUS	AY130767				
DEFINITION	Caenorhabditis elegans UDPgalNAc:GlcNAc[beta]-R				
ACCESSION	AY130767				
VERSION	AY130767.1	GI:22415754			
KEYWORDS					
SOURCE	Caenorhabditis elegans				
ORGANISM	Caenorhabditis elegans				
REFERENCE	1 (bases 1 to 1152)				
AUTHORS	Kawar,Z.S., Van Die,I. and Cummings,R.D.				
TITLE	Molecular Cloning and Enzymatic Characterization of a				
	UDP-GalNAc:GlcNAc[beta]-R beta 1,4-N-Acetylglucosaminyltransferase				
	from Caenorhabditis elegans				
JOURNAL	J. Biol. Chem. 277 (38), 34924-34932 (2002)				
PUBMED	12167666				
REFERENCE	2 (bases 1 to 1152)				
AUTHORS	Kawar,Z.S., Van Die,I. and Cummings,R.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-JUN-2002) Biochemistry and Molecular Biology,				
	University of Oklahoma Health Sciences Center, 975 NE 10th St. BRC				
	Rm. 417, Oklahoma City, OK 73104, USA				
FEATURES	Location/Qualifiers				
source	1..1152				
	/organism="Caenorhabditis elegans"				
	/mol_type="mRNA"				
	/db_xref="taxon:6239"				
CDS	1..1152				
	/note="Ce[beta]4GalNAcT, similar to Caenorhabditis elegans				
	hypothetical protein Y7357A.7"				
	/codon_start=1				
	/product="UDPgalNAc:GlcNAc[beta]-R				
	{beta}1,4-N-acetylglucosaminyltransferase"				
	/protein_id="AA095168.1"				
	/db_xref="GI:22415755"				
	/translation="MAFRHLAVARIKSLIVLCAVILLVHAMITKPSLYENITTSST				
	LIADVDAMEAVLGNASTSDLDLTWNTFSPISEVNTGTSMEDIRPLFPDNOQLQF				
	CNQTPEPHLVGPIRVFLDSDPDKTEKIKYVPTFHAGHGMPCQVARRVAIVPYDRE				

HLRLMLHNLHSLAKQOLDVAIFIVEOVANOTFNKRLMNVGVASRLYPWOCFIF
HDVDLIPEDDRNLITPCIPRHSVAIDENKYLPSALFEGISALTQHLKKNIFS
NDPMGGSDDLATRTSMAGLKVSRYPQIARVKMIKSTETATNVNCRKXKIMGQT
KRRWTDGJSLNKKYKLVNLEKPLYTRAVVDLLEKDCRRLRDPFPC"

ORIGIN

Query Match 100.0%; Score 1152; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTTTGTCAATTTGGCAGATCGCGACACTCAAGTCCTTGTCTCGTACTTTGTGCCGT 60
DB 1 ATGGCTTTTGTCAATTTGGCAGATCGCGACACTCAAGTCCTTGTCTCGTACTTTGTGCCGT 60
QY 61 CTTCATTTAGTTCATGCAATGATTTATTAAGTTCATCGCTTTACGAGAACTTACTATC 120
DB 61 CTTCATTTAGTTCATGCAATGATTTATTAAGTTCATCGCTTTACGAGAACTTACTATC 120
QY 121 GGCTCTCGACCTTATTTGCGACGTCGACGCAATGAGGCAAGTCTCGGAAATACGGCT 180
DB 121 GGCTCTCGACCTTATTTGCGACGTCGACGCAATGAGGCAAGTCTCGGAAATACGGCT 180
QY 181 TCACATTCGAGATGATCTACTTGTATAGTGAATCCAGCTTTTCCAGATTTCTGAAGT 240
DB 181 TCACATTCGAGATGATCTACTTGTATAGTGAATTCACGATTTTCCAGATTTCTGAAGT 240
QY 241 AATCAGACTAGTTTATGAGAGCAATCGCAATCGCTTCCCGCAACAGACTCTT 300
DB 241 AATCAGACTAGTTTATGAGAGCAATCGCAATCGCTTCCCGCAACAGACTCTT 300
QY 301 CAATTCGTATGACAGACACCTCCCACTCGTGGACCCATCGGTATTCCTCGATGAG 360
DB 301 CAATTCGTATGACAGACACCTCCCACTCGTGGACCCATCGGTATTCCTCGATGAG 360
QY 361 CCGGACTTCAAACTCTCGAGAAATTTATCCGACACGACCGCGGTGACATGGAATG 420
DB 361 CCGGACTTCAAACTCTCGAGAAATTTATCCGACACGACCGCGGTGACATGGAATG 420
QY 421 CCTAAGGATGTGTGCAAGCAATCGTGTCTATTTATGAGCCCTATAGATCGGAA 480
DB 421 CCTAAGGATGTGTGCAAGCAATCGTGTCTATTTATGAGCCCTATAGATCGGAA 480
QY 481 GCACTTTGAGAAATATGCTCCACAAATTTGCACTGTTGCTCGCAAAACAATTTGAC 540
DB 481 GCACTTTGAGAAATATGCTCCACAAATTTGCACTGTTGCTCGCAAAACAATTTGAC 540
QY 541 TATGCAATTTTCAATTTGAGCAAGTGGCAATCGATTCGCGGAAACTATATG 600
DB 541 TATGCAATTTTCAATTTGAGCAAGTGGCAATCGATTCGCGGAAACTATATG 600
QY 601 AAGCTTGATACGACGATGACGCTCTACCCATGGGCAAGTCTTCACTTTATGAT 660
DB 601 AAGCTTGATACGACGATGACGCTCTACCCATGGGCAAGTCTTCACTTTATGAT 660
QY 661 GTGATTTAAGTCCCGGAAGATGACCGTACCTGTACAGTGTCCAAATTCACACGTCAT 720
DB 661 GTGATTTAAGTCCCGGAAGATGACCGTACCTGTACAGTGTCCAAATTCACACGTCAT 720
QY 721 ATGAGTGTAGGATCGATTAATTAATTTCAATTTTCAATTTGCGGAGATCTTCGAGG 780
DB 721 ATGAGTGTAGGATCGATTAATTAATTTCAATTTTCAATTTGCGGAGATCTTCGAGG 780
QY 781 ATGAGTGTAGGATCGATTAATTAATTTCAATTTTCAATTTGCGGAGATCTTCGAGG 840
DB 781 ATGAGTGTAGGATCGATTAATTAATTTCAATTTTCAATTTGCGGAGATCTTCGAGG 840
QY 841 GATTGGGCGAGAGAGAGAGATTTGGCGACGAAATCGATGAGTGTGAATGAAATTTGG 900
DB 841 GATTGGGCGAGAGAGAGAGATTTGGCGACGAAATCGATGAGTGTGAATGAAATTTGG 900
QY 901 TCAAGATATCGGACAAATTTGCGATATTAATTAATTAATTAATTAATTAATTAATTAAT 960
DB 901 TCAAGATATCGGACAAATTTGCGATATTAATTAATTAATTAATTAATTAATTAATTAAT 960

QY 961 AATCCAGTTAATAATGCCGCTACAAAATTAATGGCCAAAGCAGGCGGATGACACGT 1020
DB 961 AATCCAGTTAATAATGCCGCTACAAAATTAATGGCCAAAGCAGGCGGATGACACGT 1020
QY 1021 GACGCGCTTAAGCAATCTGAATTAAGTCTGTAAATCTGGAATTTGAAGCTTCTACACT 1080
DB 1021 GACGCGCTTAAGCAATCTGAATTAAGTCTGTAAATCTGGAATTTGAAGCTTCTACACT 1080
QY 1081 CGAGCGCGTGTGATTTGCTCGAAAAGAGCTCGCCGCGGAGACTGTGAGAGGACTTCA 1140
DB 1081 CGAGCGCGTGTGATTTGCTCGAAAAGAGCTCGCCGCGGAGACTGTGAGAGGACTTCA 1140
QY 1141 ACGTGTTTTAG 1152
DB 1141 ACGTGTTTTAG 1152

RESULT 2

AY533306 1152 bp mRNA linear INV 23-FEB-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AY533306
AY533306
AY533306.1 GI:42601221
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Pelodetinae; Caenorhabditis.
1 (bases 1 to 1152)
Griffiths,J.S., Huffman,D.L., Whitacre,J.L., Barrows,B.D.,
Maroquin,L.D., Muller,R., Brown,J.R., Henmet,T., Esko,J.D. and
Arolian,R.V.

TITLE

Resistance to a bacterial toxin is mediated by removal of a
conserved glycosylation pathway required for toxin-host
interactions
J. Biol. Chem. 278 (46), 45594-45602 (2003)

JOURNAL

PUBMED
12944392
Griffiths,J.S., Huffman,D.L. and Arolian,R.V.
Direct Submission
Submitted (26-JAN-2004) Biology, Univ. California, San Diego, 9500
Gilman Dr, La Jolla, CA 92093, USA

FEATURES

source

1..1152
/organism="Caenorhabditis elegans"

/mol_type="mRNA"

/db_xref="taxon:6239"

/chromosome="I"

/map="Y73E7A"

1..1152

/gene="bre-4"

1..1152

/gene="bre-4"

/gene="bre-4"

/note="glycosyltransferase; required for susceptibility to
Bacillus thuringiensis Cry toxins"

/codon_start=1

/product="BRE-4"

/protein_id="AA521308.1"

/db_xref="GI:42601222"

/translation="MAFRHLAVARLKSLLVLCVALLVHAMTKLPSLYENTLTGSS
TLADVDMAEVLGNFTASTSDLDLTNASTFSPISFVNQTSFMDIRPIFPNQTLQ
KNQTPHLVGPTRFLDEDPFKLEKIYDTHAGGMPKDCVABRVAIIVYRDRE
AHLRLMLHNLHSLAKQOLDVAIFIVEOVANOTFNKRLMNVGVASRLYPWOCFIF
HDVDLIPEDDRNLITPCIPRHSVAIDENKYLPSALFEGISALTQHLKKNIFS
NDPMGGSDDLATRTSMAGLKVSRYPQIARVKMIKSTETATNVNCRKXKIMGQT
KRRWTDGJSLNKKYKLVNLEKPLYTRAVVDLLEKDCRRLRDPFPC"

ORIGIN

Query Match 100.0%; Score 1152; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	ANGGCTTTGGCAATTTGGGAGTGGCCGACATCAAGTCGTGGCTCGTACTTTGGGCCGT	60
Db	1	ANGGCTTTGGCAATTTGGGAGTGGCCGACAGCTCAAGTCGTGGCTCGTACTTTGGGCCGT	60
OY	61	CTTCATATAGTTCATNGCAATGATTTATATGAATTCATCGCTTACAGAACTTACTATC	120
Db	61	CTTCATATAGTTCATNGCAATGATTTATATGAATTCATCGCTTACAGAACTTACTATC	120
OY	121	GGCTTCCTGCACCTTTATTTGCCGAGTCGACGCAATGAGGAGCATGCTCGGAAATACGACT	180
Db	121	GGCTTCCTGCACCTTTATTTGCCGAGTCGACGCAATGAGGAGCATGCTCGGAAATACGACT	180
OY	181	TCCACTTCGGATNGATCTACTTGATACCGGAAATTCACAGTTTCAACGATTTCTGAAGTT	240
Db	181	TCCACTTCGGATNGATCTACTTGATACCGGAAATTCACAGTTTCAACGATTTCTGAAGTT	240
OY	241	AATCAGACTATGTTTATGAGGAGCACTTGTCCAAATCCGTCCCGGCAACCAAGACTCTT	300
Db	241	AATCAGACTATGTTTATGAGGAGCACTTGTCCAAATCCGTTCCTCCCGCAACCAAGACTCTT	300
OY	301	CAATTCGTGTAATCAGACACTTCCCACTCGTCGGAGCCATCCGTTATTCCTGCATGAG	360
Db	301	CAATTCGTGTAATCAGACACTTCCCACTCGTCGGAGCCATCCGTTATTCCTGCATGAG	360
OY	361	CCCCACTTCAAACTCTTCGAGAAAATCTATCCGACACAGCACGCGCGGTGGACATGGAATG	420
Db	361	CCCCACTTCAAACTCTTCGAGAAAATCTATCCGACACAGCACGCGCGGTGGACATGGAATG	420
OY	421	CCTAAGAGTTGTGTGGCAAGGCACTGGTTCCTATTATGTGACCCTATAGAGATCGGAA	480
Db	421	CCTAAGAGTTGTGTGGCAAGGCACTGGTTCCTATTATGTGACCCTATAGAGATCGGAA	480
OY	481	GCACATTTGGAATPATGCTCCACAATTTGGCACTCGTTCTCGCCAAACAACAATTTGGAC	540
Db	481	GCACATTTGGAATPATGCTCCACAATTTGGCACTCGTTCTCGCCAAACAACAATTTGGAC	540
OY	541	TATGCAATTTTCATTTGTGGAGCAAGTGGCGAATCAGAGCTTTAATTCGGGGAAACTAATG	600
Db	541	TATGCAATTTTCATTTGTGGAGCAAGTGGCGAATCAGAGCTTTAATTCGGGGAAACTAATG	600
OY	601	AACGTTGGATPGAGAGTGAATCAAGCCCTACCAATGGAGTGCCTTATCTTCAATGAT	660
Db	601	AACGTTGGATPGAGAGTGAATCAAGCCCTTACCAATGGAGTGCCTTATCTTCAATGAT	660
OY	661	GTCGATTTACTGCCGAGATGACCGGTAACTGTACAGGTGTCCAATTCACACACGTCAT	720
Db	661	GTCGATTTACTGCCGAGATGACCGGTAACTGTACAGGTGTCCAATTCACACACGTCAT	720
OY	721	ATGAGTGTAGCGATGATPAAATTCOAATTAATACTTCOAATTCGGCGATCTTGGCGGGA	780
Db	721	ATGAGTGTAGCGATGATPAAATTCOAATTAATACTTCOAATTCGGCGATCTTGGCGGGA	780
OY	781	ATCAGTGTGACATPACAAAGATCACTGTGAAGAAATCAATGGAATTTTGGAAATTTTTGG	840
Db	781	ATCAGTGTGACATPACAAAGATCACTGTGAAGAAATCAATGGAATTTTGGAAATTTTTGG	840
OY	841	GGTTGGGCGGAGAGAGACGATTTTGGCGACGAACATGATGAGTGTGACTGAAAGTT	900
Db	841	GGTTGGGCGGAGAGAGACGATTTTGGCGACGAACATGATGAGTGTGACTGAAAGTT	900
OY	901	TCGAAGATATCCGACACAATATTGCAGATATAAATGATTTAAGCACTGCGAGGAAAGCAGC	960
Db	901	TCGAAGATATCCGACACAATATTGCAGATATAAATGATTTAAGCACTGCGAGGAAAGCAGC	960
OY	961	AATCAGTATATTAATGCGCTGTACAAATTAATGCGCCAAAGGAGCGCGATGACACGT	1020
Db	961	AATCAGTATATTAATGCGCTGTACAAATTAATGCGCCAAAGGAGCGCGATGACACGT	1020
OY	1021	GACGGCCTTAAGCAATCTGAAGTATAGCTCTGTAATCTGGAATTTGAAGCCTCTCTACACT	1080
Db	1021	GACGGCCTTAAGCAATCTGAAGTATAGCTCTGTAATCTGGAATTTGAAGCCTCTCTACACT	1080
OY	1081	CGAGCCGTGTCGATTTTCTCGAAAAGACTGCGCGCGGAGCTTCGAGAGGGAATTTTCCA	1140

Db	1081	CGACCCGTCGATCTTCTCGAAAAAGACTGCCCGGAGCTCGAAGGCACTTTCCA	1140
Qy	1141	ACGCTTTTTAG 1152	
Db	1141	ACGCTTTTTAG 1152	
RESULT 3			
AC025727/c			
LOCUS	AC025727	63747 bp	DNA linear INV 27-MAY-2005
DEFINITION	Caenorhabditis elegans	cosmid Y73E7A, complete sequence.	
ACCESSION	AC025727		
VERSION	AC025727.3	GI:14530915	
KEYWORDS	HTG.		
ORGANISM	Caenorhabditis elegans		
SOURCE	Caenorhabditis elegans		
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;		
	Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.		
REFERENCE	1 (bases 1 to 63747)		
AUTHORS	C. elegans Sequencing Consortium		
CONSTRM	Genome sequence of the nematode C. elegans: a platform for		
TITLE	investigating biology		
JOURNAL	Science 282 (5396), 2012-2018 (1998)		
REFERENCE	9851916		
AUTHORS	2 (bases 1 to 63747)		
TITLE	Du,H. and Maupin,R.		
JOURNAL	The sequence of C. elegans cosmid Y73E7A		
REFERENCE	3 (bases 1 to 63747)		
AUTHORS	Waterston,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-MAR-2000) Genome Sequencing Center, Washington		
	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
	MO 63108, USA		
REFERENCE	4 (bases 1 to 63747)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-OCT-2000) Department of Genetics, Washington		
	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
REFERENCE	5 (bases 1 to 63747)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUN-2001) Department of Genetics, Washington		
	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.		
	Louis, MO 63110, USA		
REFERENCE	6 (bases 1 to 63747)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-NOV-2001) Department of Genetics, Washington		
	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.		
	Louis, MO 63110, USA		
REFERENCE	7 (bases 1 to 63747)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-MAY-2002) Department of Genetics, Washington		
	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.		
	Louis, MO 63110, USA		
REFERENCE	8 (bases 1 to 63747)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-NOV-2002) Department of Genetics, Washington		
	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.		
	Louis, MO 63110, USA		
REFERENCE	9 (bases 1 to 63747)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-NOV-2002) Department of Genetics, Washington		
	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.		
	Louis, MO 63110, USA		
REFERENCE	10 (bases 1 to 63747)		
AUTHORS	.		

CONSRM
TITLE WormBase Consortium
JOURNAL Direct Submission
Submitted (22-SEP-2004) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE
AUTHORS 11 (bases 1 to 63747)
TITLE WormBase Consortium
JOURNAL Direct Submission
Submitted (05-MAR-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE
AUTHORS 12 (bases 1 to 63747)
TITLE WormBase Consortium
JOURNAL Direct Submission
Submitted (27-MAY-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

COMMENT
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jee@anger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

For a graphical representation of this clone sequence and its analysis see:
http://www.wormbase.org/db/seq/sequence?name=Y73E7A;class=sequence

NEIGHBORING CLONE INFORMATION

The 5' clone is Y92H12A, 200 bp overlap; the 3' clone is Y71G12B, 4000 bp overlap.

NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program GeneIndex (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (http://worfdb.dfci.harvard.edu/), similarity to other proteins from Blastx analyses (http://blast.wustl.edu/), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. rRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES
source
1.63747
/organism="Caenorhabditis elegans"
/mol_type="Genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y73E7A"
5487..12773

CDS
/locus_tag="Y73E7A.3"
join(5487..5659,7957..8063,8778..8971,9035..9119,10020..10254,11408..11615,11668..11742,11836..11943,12669..12773)
/locus_tag="Y73E7A.3"
/standard_name="Y73E7A.3"
/note="contains similarity to Pfam domain PF06027 (Eukaryotic protein of unknown function (DUF914)); coded for by the following C. elegans cDNAs: YK1142a05.3, OSTF0045C02"

gene
/locus_tag="Y73E7A.3"
/product="Hypothetical protein Y73E7A.3"
/protein_id="AAG23391.3"
/db_xref="GI:60543518"
/db_xref="WormBase:Y73E7A.3"

CDS
/locus_tag="Y73E7A.4"
complement(join(14300..14457,15222..15495))
/locus_tag="Y73E7A.4"
/standard_name="Y73E7A.4"
/note="contains similarity to Pfam domain PF05835 (Synaphin protein); coded for by the following C. elegans cDNAs: YK1215a10.5, YK1728h08.3, OSTF0035H07"
/codon_start=1
/product="Hypothetical protein Y73E7A.4"
/protein_id="AAG23387.1"
/db_xref="GI:10800392"
/db_xref="WormBase:Y73E7A.4"
/translation="MAGFLMKQVNOQISEVTGKGMKDDGSEKTEGDEPVIARL EOEERKEKRGKENERERKQIGIDYAIKKKEGVMDFEGRIGRPRTPEETAA ENNAEDSLIGQLGTEQVEKAKWTATGAFETVKGFFPPGK"
complement(19944..20816)
/locus_tag="Y73E7A.5"
complement(join(19944..20130,20270..20541,20589..20816))
/locus_tag="Y73E7A.5"
/standard_name="Y73E7A.5"
/note="contains similarity to Pfam domain PF07735 (P-box associated)"
/codon_start=1
/product="Hypothetical protein Y73E7A.5"
/protein_id="AAG23386.1"
/db_xref="GI:10800391"
/db_xref="WormBase:Y73E7A.5"
/translation="MFOHFIKFOIIGDFRLSYOSALSHPLPRKCHPLKPLGEGSYI TTBDMKRLIKNNKNGKPTRLRSRDKIKLVNPRLDPEKRSGLTLDLKKMNC KRFEDLNTIKREINKYIRYWDGDLPLRLQLYNMHHWAGDILYLPYSKNPK NNAVSVPDITNN"
21160..26092
/locus_tag="Y73E7A.2"
join(21160..21238,21285..21638,23801..24363,25697..26092)
/locus_tag="Y73E7A.2"
/standard_name="Y73E7A.2"
/note="contains similarity to Ratrus norvegicus AATP protein (Apoptosis antagonizing transcription factor).; TR:Q9QYW0; coded for by the following C. elegans cDNAs:

Query Match 22.7%; Score 262; DB 2; Length 63747;
Best Local Similarity 100.0%; Pred. NO. 2e-68; Indels 0; Gaps 0;
Matches 262; Conservative 0; Mismatches 0

QY 1 ATGGCTTTGCGATTGGCAGTGGCCAGATCAAGTCGTTGCTGTTGCGGTT 60
DB 45925 ATGGCTTTGCGATTGGCAGTGGCCAGATCAAGTCGTTGCTGTTGCGGTT 45866

Db 1016 GGGCCCTGGTGGGCGCTTCGAGCGTGAACCTGCAATGCGCCAGCATCAGTGGCTATTGT 1075
Qy 459 TGTGCCCTTAGAGATCGTGAAGCACAATTGAGAAATAGTCCCAATTTGCACTCGTT 518
Db 1076 TGTGCCCTTCGGGATCGATAGCGCCATCTATTCTTCTTCGGCAACATCCACCATTT 1135
Qy 519 GCTCGCCAAACAACAATTGAGATGATGCAATTTTCATTGTGAGCAAGTGGCCGATCGAC 578
Db 1136 TCTGATGAAGACGCGGATGGCCATTCGCAATTTTCATTGTGAGCACAACCAAGCGAAGCC 1195
Qy 579 GTTTAAATGCGGGGAACTAATGACGTTGGATTCGACGTACGACGCTTACCCTG 638
Db 1196 CTTTAAATCGGGCGCCATGATGAACAATGGTATTGTGAGGCGCTTAAAGCTGTACCAAGTG 1255
Qy 639 GGAAGCTTCATCTTCATGATGATGATGATTTTACGCGCCGAAGATGACCGTAACTGTAC 698
Db 1256 GGAATGTTTAAATTCACAGATGTCATCTTCTGCTTTGGACGCGCAATCTTTTACAA 1315
Qy 699 GTGTCCAAATTCACAACGATCATATGATGATGATGATGATGATGATGATGATGATGATG 758
Db 1316 CTGTCCACGTCACGCGCGGACACATGTCAGTGCCTATAGACAGCTGAACTTCAGTTGCC 1375
Qy 759 AATATTCGGGATCTTTCGGCGGAAATCAGTGCATTAACAAAAGATCCTGAAGAAATCAA 818
Db 1376 TTAATCATCAATATTTGAGAGTGTTCCTCCGAATGACGCGGTGACACTTTCAGGCGCTTAA 1435
Qy 819 TGGATTTTGAATGATTTTGGGTTGGGGGAGAGAGACGATTTGGCGAGAAC 878
Db 1436 TGGGTTCTCAACTCGTTCTTTGGCGGGGCGGAGAGATGACGACATGTCACACAGTTT 1495
Qy 879 ATGATGCTGCTGACTGAAGTTTCAAGATATCCGACACAATTCAGATGATATTAATGAT 938
Db 1496 GAAGCAGCGAACAATTATCATATCAAGATATCCGGTCAATGCTGCTACAGATGCT 1555
Qy 939 TAAAGCACTGACGGAAGCGACGAATCC 965
Db 1556 GAAGCATCAGAAAGAAAGCCCAATCC 1582

RESULT 6
AY601103 1269 bp mRNA linear INV 02-AUG-2004
LOCUS Trichoplusia ni beta 1,4-N-acetylglucosaminyltransferase mRNA,
DEFINITION complete cds.
ACCESSION AY601103
VERSION AY601103.1 GI:47156062
KEYWORDS Trichoplusia ni (cabbage looper)
SOURCE Trichoplusia ni
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Plusiinae; Trichoplusia.
REFERENCE
AUTHORS Vadate,N. and Jarvis,D.L.
TITLE Molecular Cloning and Functional Characterization of a Lepidopteran
Insect [beta]4-N-Acetylglucosaminyltransferase with Broad
Substrate Specificity, a Functional Role in Glycoprotein
Biosynthesis, and a Potential Functional Role in Glycolipid
Biosynthesis
J. Biol. Chem. 279 (32), 33501-33518 (2004)
PUBMED 15173167
REFERENCE 2 (bases 1 to 1269)
AUTHORS Vadate,N. and Jarvis,D.L.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2004) Molecular Biology, University of Wyoming,
Laramie, WY 82071, USA
FEATURES
source
1. 1269
/organism="Trichoplusia ni"
/mol_type="mRNA"
/db_xref="taxon:7111"
1. 1266

/note="b4GalNCT; glycoeyltransferase"
/codon_start=1
/product="beta 1,4-N-acetylglucosaminyltransferase"
/protein_id="AAT11926.1"
/db_xref="GI:47156063"
/translation="MGRAATRAALRLLLVLAALAVELFSGSIIDASPLRFLYLYPLV
NATOPTLNLVERLANWPKLPSNTIESEYSINISLNHTTRASVHPSSITET
ASKDKMNTIDGAPAMISPTPLITKLMDSIKSVTTEDGVKAIEAVVTLPLCDMP
PDIGITLNTLELEWEEKFPEVEMGGRYSPPCTARHRAIIVPRDQHLAIF
LNHMRFLMKQOIEYGI FIVEQGNKDNRAKLVNMGVESQKLVASGMQCFPHID
LPLDRLVYSGCPQPRHMSASIDKLHRLPYEDI FGVSAVMTLEQFTRVNSKYM
GHEGDDMSYRTRIKINHYIARYKMSIARYMLDHKSITPNPKRYQLSQTSTQKD
GLSTLELVQVVOYHLYTHLVNIDERS"

ORIGIN
Query Match 16.7%; Score 192; DB 2; Length 1269;
Best Local Similarity 58.3%; Pred. No. 4.4e-47;
Matches 358; Conservative 0; Mismatches 250; Indels 6; Gaps 1;
Qy 358 GAGCCCGACTTCGAAACTCTCGAGAAATCTATCCGACACGACCGCGGTGACATGGA 417
Db 514 GAACTGGAATCGGAATGGGTGAGAAAAAGTTCCGAGAGTGAAGTGGGAGGCGGTAC 573
Qy 418 ATGCCTAAGATTTGTTGCAAGCATGCTGTTGCTATTATTTGTCCTTATAGATGCT 477
Db 574 TCGCGCCCACTGACGCGCCGCAATGAGTGCCTATTATGTGCATACAGATGCT 633
Qy 478 GAAGCATTGAGATTAATATGCTCCAAATTTGCACTGCTGCTGCCCAACAACATTTG 537
Db 634 CAGCAGCAATTTAGCATATCTCTGAATCATATGCAATCCGTTTGTATGAAACAGCAAT 693
Qy 538 GACTATGCAATTTCAATTTGAGAGCAAGTGGCAATTCAGACGTTTAAATCGCGGAACTA 597
Db 694 GAATATGAAATTTTCAATGTCGACGACAAAGAAACAAGAACTTCAACCGCCAAAGCTG 753
Qy 598 ATGAAGTTGATATGAGCTAGCATCAGCCCTTACCCA-----TGGAGTGGCTTATC 651
Db 754 ATGAAGTTGGCTTGTGAGAGAGCAAGAGCTGTGGCGAAGGCTGGAGTGGCTTGTG 813
Qy 652 TTTTCATGATGCTGATTTATGCCCCGGAATGACCGTAACTGTACAGTGTCCAAATTCAA 711
Db 814 TTCCACGACATGACCTTGTGCTCTCCGACACAGGAACCTGTATCTTCCGACAGACAG 873
Qy 712 CCACGTCAATATGAGTACGATCGATGATTAATTCATTAATTAATTCATTAATTTGGCGATC 771
Db 874 CCGCGACATATGTCGGCTTCCATCGATTAATTAATTCATTAATTCATTAATTCATTAAT 933
Qy 772 TTGCGCGGAATCAGTGCATTAACAAAAGTCACTGAAAGAAATCAATGATTTTGGAT 831
Db 934 TTGCGCGGAGTCTTCCGATGACCTTGAACAGTTTCACAGAGGTCAACGCGCTTCTGAA 993
Qy 832 GATTTTGGGGTTGGGGCGGAGAGCAGCATTTTGGCAGCAGAACTGATGCTGGA 891
Db 994 AAGTACTGGGGTTGGGGCGGGAAGATGATGACATCTCATATAGGTTAAAGAAATTAAT 1053
Qy 892 CTGAAAGTTTCAAGATATTCGACACAAATTCAGATATTAATGTTAAGCACTGACG 951
Db 1054 TATCATATAGCAAGATACAAAATGTCTATAGCGCGGTACGTAATGTTAGATCAAGAAA 1113
Qy 952 GAAGCAGCAATTC 965
Db 1114 TCTACCTTAATTC 1127

RESULT 7
AR506512 1189 bp DNA linear PAT 22-SEP-2004
LOCUS AR506512
DEFINITION Sequence 11472 from patent US 6703491.
ACCESSION AR506512
VERSION AR506512.1 GI:52441987
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE	Unclassified.
1 (bases 1 to 1189)	
AUTHORS	Homburger, S.A., Ebers, A.J. Jr., Erickson, C.S., Francis-Lang, H.L., Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchman, A.R.
TITLE	Drosophila sequences
JOURNAL	Patent: US 6703491-A 11472 09-MAR-2004; Exelixis, Inc.; South San Francisco, CA
FEATURES	location/Qualifiers
Source	1. 1189
ORIGIN	/organism="unknown" /mol_type="genomic DNA"
Query Match	16.6%; Score 191; DB 6; Length 1189;
Best Local Similarity	58.6%; Pred. No. 8.9e-47;
Matches 332; Conservative	0; Mismatches 235; Indels 0; Gaps 0;
Qy	399 GCACGCCGGTGAACATGGAATGCCAAGATGTTGTTCCAGAGGACGCGTGTGATATAT
Db	508 GCGCCCTGGTGGCGCTTCGAGCGCTGAAGACTGCATCCAGACGACGCGGCTATGTT
Qy	459 TGTGCCCTATAGAGATCGTGAAGCAGATTTGAGATTAATGCTCCAGATTTGCACTCGT
Db	568 TGTGCCCTCGCGCATGATAGGCCCATCTATTACTTTTCTGGCGCAATCCACCATT
Qy	519 GCTGCCCAACACACATTTGACTATGCAATTTTCTTTGAGGCAAGTGGCGAATCAGAC
Db	628 TCTATGAGAGAGGCGCATGCCCTATCCGATTTTCAATGTAGAGCAGACCAACGGGAGCC
Qy	579 GTTTAATCGCGGGAACTAATGAACGTTGATAGACGCTAGCATCAGCCCTTACCAGT
Db	688 CTTCATAGGGCTGCGCATGATGAACATGTTATTTGAGGCTTAAAGCTGTACAGTG
Qy	639 GCAGTGCCTTCATCTTTCATGATGTCATTTACTGCCCCAGATGACCTGTACAC
Db	748 GGATTTGTTTATATTCACAGATGTGATCTTCTGCTTTGAGACGCGCAATCTTACA
Qy	699 GTGTCCATTTCAACACGCTCATATGAGTGTAGCGATGATTAATCAATTATTAACCTTCC
Db	808 CTGTCCAGTCAAGCGCGCAGACACATGTCAAGTGTATAGACAGCTGAACCTTCAGGTTGCC
Qy	759 ATATTGGCGGATCTTCGGCGGATCAAGCAGCTAACAAAGATACCCGGAAGAAATGAA
Db	868 TTATTCGATCAATATTGGAGGTGTGTTCCGCAATGACGGGTAGACCTTTCAGGCCGTAA
Qy	819 TGGATTTTTCGATGATATTTTGGGGGTGGGGCGAGAGACGACATTTGGCGACGAGAAC
Db	928 TGGGTTCTCAAACTGCTTTCTTTGGCTGGGGCGGCGAGATGACGATGTCCACAGGTT
Qy	879 ATCGATGCTGGAATGAAGTTTCAAGATATCCGACACAAATTCGACGATTAATAATGAT
Db	988 GAAAGAGCGCAACCTATTCATATACAGGTATCCGGTCAACATAGCCGCTCAAGATGCT
Qy	939 TAAAGCATCGACGGAAGCAGCAATCC
Db	1048 GAAAGCATCGAAGGAAAGGCCCAATCC
RESULT 8	
LOCUS	AY095531 1879 bp mRNA linear INV 17-APR-2002
DEFINITION	Drosophila melanogaster RE56531 full insert cDNA.
ACCESSION	AY095531
VERSION	AY095531.1 GI:20177118
KEYWORDS	FLI CDNA.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster
REFERENCE	Buxarjota, Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 1879) Stapleton, M., Brokstein, P., Hong, L., Agrayani, A., Carlson, J., Champe, M., Chavez, C., Doree, V., Dresnek, D., Fairlan, D., Frise, E.,

ORIGIN	FEATURES	COMMENT
Query Match: 16.6%; Score 191; DB 2; Length 1879;		
Best Local Similarity: 58.6%; Pred. No. 9,4e-47;		
Matches 332; Conservative 0; Mismatches 235; Indels 0; Gaps 0;		
399 GCACGCCGCTGAGCATGAATGCTTAAGAATTGTGTCGAAGGATCGTGTGCTATTAT 458		
Db 630 GCGCCCTGTGGGCGCCCTTCGAGCGCTGAACCTGCAAGCCAGCATCAGCGGTATTATGT 689		
459 TGTGSCCTATAGAGATCGTAGAGACATTGGAAGTAATGCTCCACAATTGGACCTGGTT 518		
Db 690 TGTGSCCTTCGCGATCGATGATCGCCATCTATTACTTTCTGTGCAACATCCACCAATT 749		
519 GCTCGCCAAACAACATTTGACTATGCAATTTTCATTTGTGAGAGCACTGGCGATCAAC 578		
Db 750 TGTGATAGACAGGAGCATGCGCTATGCAATTTTCATTTGTAGAGAGACAAACGGAGACC 809		
579 GTTTAATCGCGGGAACCTAATGAACGTTGGATAGACGATGACATCAGCGCTTACCATG 638		

Db 810 CTTTAACTGGGCGGCATGATGAACATGGTTATTGGAGGCGCTTAAGCTGTACAGTG 869
Qy 639 GCAAGTCTTCATCTTTCATGATGATGCATTTACTGCCCGAAGATGACCTGTACAC 698
Db 870 GGATTTGTTTAAATTCACAGATGTCATCTTCGCTTTGGAGACCGCAATCTTTACAA 929
Qy 699 GTGTCCATTTCAACACAGTCATATGATGATGATGATGATGATGATGATGATGATGAT 758
Db 930 GTGTCCACGTGACCGCGGACACATGTCATGATGATGATGATGATGATGATGATGATGAT 989
Qy 759 ATATTGCGGATCTTGGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 818
Db 990 TTAATCATCAATTTTGAAGTGTTCGCAATGACCGGTGACATTTTCAGGCGGTAA 1049
Qy 819 TGGATTTTGAATGATTTTGGGTTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
Db 1050 TGGGTTCTCAAACTGCTTCTTTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1109
Qy 879 ATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938
Db 1110 GAAGCAGCGCAACCTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1169
Qy 939 TAAACATCGACGAG 965
Db 1170 GAAGCATCAAGAAAG 1196

RESULT 9
AB036858 2796 bp mRNA linear INV 15-JUL-2000
LOCUS Ciona intestinalis Ci-b4Gal-T mRNA for beta 4
DEFINITION galactosyltransferase, complete cds.

AB036858.1 GI:9229931
AB036858 Ci-b4Gal-T; beta 4 galactosyltransferase.
VERSION
KEYWORDS
SOURCE Ciona intestinalis
ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

REFERENCE
AUTHORS 1 (sites)
TITLE Hotta,K., Takahashi,H., Asakura,T., Satoh,B., Takatori,N.,
Sato,Y. and Satoh,N.
JOURNAL Characterization of Brachyury downstream notochord genes in the
AUTHORS Ciona intestinalis embryo
TITLE Unpublished

JOURNAL 2 (bases 1 to 2796)
REFERENCE Hotta,K., Takahashi,H. and Satoh,N.
AUTHORS Direct Submission
TITLE Submitted (13-JAN-2000) Kohji Hotta, Graduate School of Science,
JOURNAL Kyoto University, Department of Zoology; Sakyo-ku, Kyoto, Kyoto
606-8501, Japan (E-mail:kozi@acidian.zool.kyoto-u.ac.jp,
Tel:+81-75-753-4095, Fax:+81-75-705-1113)
FEATURES
source Location/Qualifiers

1. .2796
/organism="Ciona intestinalis"
/mol_type="mRNA"
/cell_type="notochord"
/note="this gene is expressed in the notochord cells"
1. .2796
/gene="Ci-b4Gal-T"
37. .1278
/gene="Ci-b4Gal-T"
/codon_start=1
/product="beta 4 galactosyltransferase"
/protein_id="BAB00635.1"
/db_xref="GI:9229931"
/translation="MSLRHYNGLVRRRTNGRAKWLTLCLILPAELIVYVPRH
DVLIVSLHNTKIDWIRGRTNCEAFGRNSLKDSALCPVHKAIIIVPHNRSHLRQ
LKGDELGENDESDLDLVKNYTLVPGKFRPAQCVPVHKAIIIVPHNRSHLRQ
FLKAIRHVMGRQADYGVFVHOSGTGTFKAKILNIGYEALIKQDDYDFIFHDVLL
LAEDDRNLVRCADVPRIHLSVIGIDKMDYQLPYDALFGVIAMTYDQFAQVNGVSNBYWG

WGAEDDDMYRIRILHSLCLGERAQYDVARYRNAHYHSDKSNVNPYRLLVGAEROR
HGLNNHLYSLVETQLPLATYNTISADVGPPLNPATYVGFEGIDILVVISLILLMLTM
TCCTCVKRSLPHVLLCPRRIP"

Query Match 16.2%; Score 186.8; DB 2; Length 2796;
Best Local Similarity 53.6%; Pred. No. 1.9e-45;
Matches 389; Conservative 0; Mismatches 337; Indels 0; Gaps 0;

Qy 354 CGATAGAGCCGACCTTCAAACTCTGAGAAATCTATCGGACAGCGCGGTGAGCA 413
Db 375 CGAGAGCGTTAGCTTGACAGATCTAGTAAACCTACAGTTGGTTGACCTGTGGGGG 434
Qy 414 TGGATGCTTAAAGATTTGTTGCAAGGATCGTGTGTATATTTGCTTATAGAGA 473
Db 435 GTTCAAAACCGGCGCATATGTCTTCGTCACAAAGGTGGATTAATGCTCCGACCGTAA 494
Qy 474 TGTGAAGCACATTTGAGATTAATGCTCCACATTTGCACTGCTGTGCGCAACACAA 533
Db 495 CCGTAGAGAGACACTACCGCAGTTTCTAAAGCTATACCTCTGTTATGCAAGCGGACAGA 554
Qy 534 ATTGACTATGCAATTTTCATTTGTGAGAGAGATGGCGAATCAGAGTTTATCGGGGAA 593
Db 555 AACCCAGCTATGATGTTTCTGTTGTCACCAATCCGACACCGGACATTTAACAGGCAAA 614
Qy 594 ACTAATGAACGTTGATACGACGTATCAGCCTCTACCCATGAGAGAGAGAGAGAGAGAG 653
Db 615 GCTGTTAAATATCGGCTACACTGAAAGCGGTGAAAGAGATGATTAATGCTTATCTT 674
Qy 654 TCAATGATGCTGATTTATCTGCGGAAAGATGACCTGTATACAGTGTCAATTCACAC 713
Db 675 TCACAGATGCGACCTGCTGCTGAGAGACGACCGCATGTGATGATGATGATGATGATGAT 734
Qy 714 AGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
Db 735 CAGACATTTATCTGTTGTTGATGACAAATGGATTTTCCAACTGCTATGACGACTGTT 794
Qy 774 CCGCGGAATCAGTGCATTAACAAAGATCAGCTGAGAAATCAATGATTTTGCAGTGA 833
Db 795 CGGTGAGATTATTTGCAATGACCAAGAGCAAAATTTGCGGAGTATTCAGAGCA 854
Qy 834 TTTTGGGGTGGGGGCGGAG 893
Db 855 ATACTGGGATGAGGGGTGCGAAGATGACGACATGATGAGAGATTTCTGATAGCTGTCT 914
Qy 894 GAAAGTTCAAGATATCCGACACAAATTGACAGATTAATGATTAAGACATCGACGGA 953
Db 915 TGGTTTGAAGAGGACAAATATGACGTTGACAGATGACAGATGAGGCTTACACCTTCGGA 974
Qy 954 AGCGAGATTCAGATTAATTAATGCGCTACCAAAATATGAGGCAACGAGCGCGGATG 1013
Db 975 CAAGACCAACGAGTAAATCCATATGTTTACACCTTACCTGTTGGGCGCGGAGAGACA 1034
Qy 1014 GACACGTGACGGCTTACAGCAATCTGAAGTATAGCTCGTAATTCGGAATTTGAGCTCT 1073
Db 1035 GAGGCAATGATGATTAACCAACCTTCATACAGCTTGTAGAGAAAACACAGTACAT 1094
Qy 1074 CTACAC 1079
Db 1095 ATACAC 1100

RESULT 10
AK116794 2082 bp mRNA linear INV 30-NOV-2002
LOCUS Ciona intestinalis cDNA, clone:cieg035d14, full insert sequence.
DEFINITION AK116794
ACCESSION AK116794.1 GI:23589473
VERSION AK116794.1
KEYWORDS F11 cDNA.
SOURCE Ciona intestinalis
ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

CDS

/note="synonym: B4galc3"
 /db_xref="GeneID:494342"
 276...1463
 /gene="MGCC72586"
 /codon_start=1
 /product="similar to UDP-Gal:betaGalCNC beta
 1,4-galactosyltransferase, polypeptide 3"
 /protein_id="AAH61812.1"
 /db_xref="GI:38197704"
 /db_xref="GeneID:494342"
 /translation="MLRLERPCITALLVGSQAVMMYLSIGFRSLALRGDRDP
 TEDYSHPDVYNSLHPGAPGAPLQVLPDCERSPLVGPVSFSPVSLAE
 IYERNRVSQGRYPAGEPRSRATIIIVPHRABRHLLYLHLPLROOLAYGI
 YVHQNGTFRNKLINVGREALRDEMDCLFLHVDLLPENDINGLVCDPRGRH
 VVAVMNRFGYSLPYQYFGVSALTPOYLKMGFENYMGWGGDDDIATYVRLAGM
 KISRPTSGVGHKMGVGRDKNENRPHRFDLLVTRQNSWTDGNSLTYRLARLGL
 PLTNITADIIGTDPGRAPSGPRYPSSQAFOEMLQRRPARPGPLPYANHTAP
 GSH"

ORIGIN

Query Match 14.6%; Score 168.4; DB 9; Length 2007;
 Best Local Similarity 53.4%; Pred. No. 8.6e-40;
 Matches 377; Conservative 0; Mismatches 326; Indels 3; Gaps 1;

377 TCGAGAAATCTATCCGACACGCGCGGTGACATGATGCTTAAGATTGTGTG 436
 |||||
 583 TCGTGAAGGGAGATCCCGGGTGGATCAAGGGCGCGTACCGTCTGACAGATGTGAC 642
 |||||
 437 CAAGCATGTGTGCTATTTATTTGCTCTATAGAGATGCTGAACCATTTGAGATTA 496
 |||||
 643 CTGCGTCCCGACAGCATATGTCGCCACCGTCCCGGAGACACACTCCGACTGC 702
 |||||
 497 TGTCCAAATTTGCACTGCTGCTGCGCAACAACAATTTGATGCAATTTTCAATG 556
 |||||
 703 TACTCTATACCTGCAACCCCTCTCAGCGCAGACGCTGCTTATGACATCTATGTCA 762
 |||||
 557 TCGAGCAAGTGGCGAATCAGACCTTTAATCGCGGAACTAATGAACGTTGATACGAG 616
 |||||
 763 TCACACAGGTGGAATGGAACGTTTACAGGGCCAAACGTCGAAAGTAAAGGCTCAG 822
 |||||
 617 TAGCATCAGGCTCTTACCCATGCGAGTCTTCACTTTTCAATGATGTCATTAAGTCC 676
 |||||
 823 AAGCCCTGCTGACGAAGAGTGGGACGCTGCTGTTTACACGATGTGACCTCTCCAG 882
 |||||
 677 AAGATACCGTACCTGTACAGCTGT---CAATTAACACAGCTCATATGATGTATGCA 733
 |||||
 883 AAAACGACATACCTGTATGTGTGACCCCGGGGACCCCGCAGTGTGCTGCCA 942
 |||||
 734 TCGATTAATTCATTAATACTTCATATTGCGCGATCTTGGCGGATACATGCACTAA 793
 |||||
 943 TGAACAAGTTTGGATACAGCTTCCGTAACCCCAAGTACTTTGGCGGAGTTTCAAGCA 1002
 |||||
 794 CAAGAATCACTGMAAATAATCATGATTTTGAATGATTTTGGGGTGGGGCGAG 853
 |||||
 1003 CTCCTACCAAGTACCTGAAGATGAATGGCTTCCCAATGAATCTGGGGCTGGGGCG 1062
 |||||
 854 AGGACACGATTTGGGAGAGAAACATTCATGCTGCACTGAAGTTCAGATATCCCA 913
 |||||
 1063 AGGATACGACATGTGTACCAAGGCTTCCGCTGCTGGAAGATTTTCCGACCA 1122
 |||||
 914 CACAATTCGACGATTAATAATGATTAAGACATCGACGAGGAGCAAGATCCAGTTAATA 973
 |||||
 1123 CCTCTGTGGGACATACAGATGTGTAGAGACAGAGGGGATAAAGAAATGAGGAAACC 1182
 |||||
 974 AATGCGGCTACAAATTAATGAGCCAAACGAGCGCGATGACACGTCGCGCTTAAGCA 1033
 |||||
 1183 CCCACAGATTGTGACCTCCGCTGCGTACCCAGATTTCTTGACACAAAGATGAATAACT 1242
 |||||
 1034 ATCTGAAGTAAAGCTGTGAATCTGGAATTTGAAGCTCTCTACAC 1079
 |||||
 1243 CACTTAACATACGACGTGCTGGCAAGAGCTGGGCTCTCTACAC 1288
 |||||

RESULT 15
 BT020916
 LOCUS
 DEFINITION
 Bos taurus UDP-Gal:betaGalCNC beta 1,4-galactosyltransferase,
 polypeptide 3 (B4GALT3), mRNA, complete cds.

ACCESSION
 BT020916
 VERSION
 FLI CDNA.
 SOURCE
 Bos taurus (cow)

REFERENCE
 Smith,T.P., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 1 (bases 1 to 1921)
 Caasas,E., Wray,J.E., White,J., Chu,J., Fahrnenkrug,S.C.,
 Bennett,C.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A.,
 Chitko-McKown,C.G., Petrea,G., Holt,I., Karamcheva,S., Liang,F.,
 Quackenbush,J. and Keefe,J.W.
 Sequence evaluation of four pooled-tissue normalized bovine CDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

COMMENT
 FEATURES
 source

gene

CDS

/note="B4GALT3"
 /note="putative, similar to GeneID: 8703"
 215...1405
 /gene="B4GALT3"
 /codon_start=1
 /product="UDP-Gal:betaGalCNC beta 1,4-
 galactosyltransferase 3"
 /protein_id="AAH08933.1"
 /db_xref="GI:59858197"
 /translation="MLRLERPCITALLVGSQAVMMYLSIGFRSLALRGDRDP
 AFDYSHPDVYNSLHPGAPGAPLQVLPDCERSPLVGPVSFSPVSLAE
 IYERNRVSQGRYPAGEPRSRATIIIVPHRABRHLLYLHLPLROOLAYGI
 YVHQNGTFRNKLINVGREALRDEMDCLFLHVDLLPENDINGLVCDPRGRH
 VVAVMNRFGYSLPYQYFGVSALTPOYLKMGFENYMGWGGDDDIATYVRLAGM
 KISRPTSGVGHKMGVGRDKNENRPHRFDLLVTRQNSWTDGNSLTYRLARLGL
 PLTNITADIIGTDPGRAPSGPRYPSSQAFOEMLQRRPARPGPLPYANHTAP
 GSH"

ORIGIN

Query Match 14.6%; Score 168.2; DB 4; Length 1921;
 Best Local Similarity 52.3%; Pred. No. 9.8e-40;
 Matches 421; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, (using frame_plus.p2n model)

Run on: January 9, 2006, 00:31:56 ; Search time 6343 Seconds
(without alignments)
3432.294 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCV.....VLLKDKRRRLRDPFPCF 383

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 segs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US10661430/runac_06012006_123134_6236/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10661430@cgn2.1 1.4939 @runac_06012006_123134_6236 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEBUFFER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	1152	2	AY130767 Caenorhab
2	2031	100.0	1152	2	AY533306 Caenorhab
3	1275	62.8	63747	2	AC025727 Caenorhab

4	775.5	38.2	1269	2	AY601103	AY601103 Trichoplui
5	741	36.5	2265	5	BC077601	BC077601 Xenopus t
6	740	36.4	2098	5	CR942597	CR942597 Xenopus t
7	736.5	36.3	2681	5	BC068719	BC068719 Xenopus t
8	731.5	36.0	2082	2	AK116794	AK116794 Clona Int
9	730	35.9	2626	5	BC041742	BC041742 Xenopus t
10	729	35.9	2340	5	BC075452	BC075452 Xenopus t
11	725	35.7	1952	9	AF142671	AF142671 Mus muscu
12	725	35.7	1962	9	BC013619	BC013619 Mus muscu
13	721	35.5	1927	4	BT020682	BT020682 Bos tauru
14	720	35.5	1329	5	BC087436	BC087436 Xenopus t
15	720	35.5	1921	4	BT020916	BT020916 Bos tauru
16	720	35.5	1949	4	BT020868	BT020868 Bos tauru
17	717	35.3	1182	8	HSUDPCAL	Y12509 Homo sapien
18	717	35.3	1193	8	AB024435	AB024435 Homo sapi
19	717	35.3	1338	9	AY117537	AY117537 Cricetulu
20	717	35.3	1912	8	BC009985	BC009985 Homo sapi
21	717	35.3	1915	8	BC000276	BC000276 Homo sapi
22	717	35.3	1934	8	BC006099	BC006099 Homo sapi
23	717	35.3	1938	6	BD206381	BD206381 Human nuc
24	717	35.3	1938	6	AR400696	AR400696 Sequence
25	717	35.3	1938	6	AX013200	AX013200 Sequence
26	716	35.3	2007	9	BC061812	BC061812 Rattus no
27	714	35.2	1885	8	AK092493	AK092493 Homo sapi
28	709	34.9	1839	6	CO590796	CO590796 Sequence
29	709	34.9	1879	2	AY095531	AY095531 Drosophi
30	709	34.9	2269	2	AF132158	AF132158 Drosophi
31	706	34.8	1920	8	AF038661	AF038661 Homo sapi
32	704	34.7	1189	6	AK506512	AK506512 Sequence
33	694	34.2	1912	8	AF020921	AF020921 Homo sapi
34	684.5	33.7	1954	5	CR760293	CR760293 Xenopus t
35	684	33.7	1853	6	BD191505	BD191505 Secreted
36	680.5	33.5	2279	5	GGU19890	U19890 Gallus gall
37	675	33.2	4093	6	CO590795	CO590795 Sequence
38	675	33.2	35217	14	AC017131	AC017131 Drosophi
39	675	33.2	165177	2	AC007588	AC007588 Drosophi
40	675	33.2	180425	2	AC008229	AC008229 Drosophi
41	675	33.2	229685	2	AE003815	AE003815 Drosophi
42	671.5	33.1	1262	9	AF118896	AF118896 Cricetulu
43	670	33.0	2420	5	GGU19889	U19889 Gallus gall
44	668.5	32.9	2304	6	A38781	A38781 Sequence 7
45	668.5	32.9	2304	6	I49897	I49897 Sequence 7

ALIGNMENTS

RESULT 1
AY130767 1152 bp mRNA linear INV 17-SEP-2002
LOCUS
DEFINITION Caenorhabditis elegans UDPgalNAc:GlcNAc(beta)-R
{beta}1,4-N-acetylgalactosaminyltransferase mRNA, complete cds.
ACCESSION AY130767
VERSION AY130767.1 GI:22415754
KEYWORDS
SOURCE
ORGANISM

Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Pelodietidae; Caenorhabditis.
REFERENCE
AUTHORS Kawai, Z.S., Van Die, I. and Cummings, R.D.
TITLE Molecular Cloning and Enzymatic Characterization of a
UDP-GalNAc:GlcNAc-beta-1,4-N-Acetylglactosaminyltransferase
from Caenorhabditis elegans
J. Biol. Chem. 277 (38), 34924-34932 (2002)

JOURNAL PUBLISHED
REFERENCE Kawai, Z.S., Van Die, I. and Cummings, R.D.
AUTHORS Direct Submission
TITLE Submitted (10-JUN-2002) Biochemistry and Molecular Biology,
JOURNAL University of Oklahoma Health Sciences Center, 975 NE 10th St. BRC
Rm. 417, Oklahoma City, OK 73104, USA
FEATURES
source 1. 1152

CDs
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
1..1152
/note="Ce(beta)4GalNAC; similar to Caenorhabditis elegans
hypothetical protein Y73E7A.7"
/product="UDPGalNAC:GlcNAc(beta)-R
(beta)1,4-N-acetylglucosaminyltransferase"
/protein_id="AAM95168.1"
/db_xref="GI:22415755"

ORIGIN

Alignment Scores:
Pred. No.: 2 84e-179 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-661-430-1 (1-383) x AY130767 (1-1152)

QY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuSerLeuLeuValLeuCySalAlaVal 20
Db 1 ATGGCTTTTCGTCACTTGGCAGTGCAGACCTCAAGTCGTTGCTGCTGCTGCTGCTGCT 60
QY 21 LeuLeuLeuValHisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThrIle 40
Db 61 CTTCATTTAGTCAATGCAATGATTATTAAGATTCATCGCTTTCGAAACCTTACATATC 120
QY 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla 60
Db 121 GGCTCTCGACCTTATTTGCCAGCTGCAGCAATGGAGGAGCTGCTCGGGAATACGGCT 180
QY 61 SerThrSerAspAspLeuAspThrTrpAsnSerThrPheSerProIleSerGluVal 80
Db 181 TCCACCTTCGAGTATGACTTGAATGAGTGAATTCAGGTTTTCACGATTTTCGAGATT 240
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
Db 241 AATCAGACTAGTTTATGAGGACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 101 GlnPheCysAsnGlnThrProProHisIleValGlyProIleArgValPheLeuAspGlu 120
Db 301 CAATTCCTGATATCGACACCTCCCGACCTGTCGAGCCCATCTGTAATTCCTCGATGAG 360
QY 121 ProAspPheLeuThrLeuGluLysIleTyrProAspThrHisAlaGlyHisGlyMet 140
Db 361 CCGGACTTCAAAACTCTCGAAGAAATCTATCCGACACGCGCGGAGGACATGGAATG 420
QY 141 ProLysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGlu 160
Db 421 CCTAAGAGATTGTGTGCAAGGCACTGTTGCTTATTTGTCCTTATGAGATCGTGA 480
QY 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnGlnLeuAsp 180
Db 481 GCACATTGGAATAATGCTCACAAATTGCACTGTTGCTGCGCAACCAATGATGAC 540
QY 181 TyrAlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMet 200
Db 541 TATGCAATTTTTCATTTGCGAGCAAGTGGCAATCAGAGCTTATTCGCGGAAACTAATG 600
QY 201 AsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAsp 220
Db 601 AACGTTGATACGACGATGATCAGGCTTACCAATGCGAGTCTTCACTTTCATGAT 660

QY 221 ValAspLeuLeuProGlnAspAspArgAsnLeuTyrThrCysProIleGlnProArgHis 240
Db 661 GTCGATTACTGCGCGAAGATGACCGCTAACCTGTACAGCTGTCCATTCACACAGCTCAT 720
QY 241 MetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGly 260
Db 721 ATGAGGTGAGCGCATCATTAATTCATTTAAACCTTCATATTCGCGCATCTTCGCGGA 780
QY 261 IleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTrp 280
Db 781 ATCAGTGCATCAACAAATGATCCCTGAAGAAATTAATGATTTTCGAAATGATTTTGG 840
QY 281 GlyTyrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuVal 300
Db 841 GGTGGGGGGAGAGAGACACAGATTTGGGACGAGAAATCATGATGCTGCTGAGTGAAGTT 900
QY 301 SerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThr 320
Db 901 TCAGATATATCCGACACAAATTCGACGATTAATTAATGATTAAGCATCGAGAACGACG 960
QY 321 AspProValAspLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTrpThrArg 340
Db 961 AATCCAGTTAATAATGCGCTACAAATTAATGGCCAAACGAAAGCCGATGGACACGT 1020
QY 341 AspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThr 360
Db 1021 GACGCGCTAAGCATTCGAAGTATTAAGCTCGTAAATCTGGAATTTGAAGCTCTTACACT 1080
QY 361 ArgAlaValAlaAspLeuLeuGlyLysAspCysArgArgGluLeuLysArgAspPhePro 380
Db 1081 CGAGCCGTCGTCGATTTGCTCGAAAAAGACTGCCCGGAGAGCTGCGAAGGAGCTTTCCA 1140
QY 381 ThrCysPhe 383
Db 1141 ACGTGTATT 1149

RESULT 2
AY533306 1152 bp mRNA linear INV 23-FEB-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS
Griffiths,J.S., Huffman,D.L., Whitacre,J.L., Barrows,B.D.,
Maroquin,L.D., Muller,R., Brown,J.R., Henne,T., Bako,J.D. and
Aroian,R.V.
Resistance to a bacterial toxin is mediated by removal of a
conserved glycosylation pathway required for toxin-host
interactions
J. Biol. Chem. 278 (46), 45594-45602 (2003)
12944392
2 (bases 1 to 1152)
REFERENCES
AUTHORS
Griffiths,J.S., Huffman,D.L. and Aroian,R.V.
Direct Submission
Submitted (26-JAN-2004) Biology, Univ. California, San Diego, 9500
Gilman Dr, La Jolla, CA 92093, USA
location/Qualifiers
1..1152
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/chromosome="I"
/map="Y73E7A"
1..1152
/gene="bre-4"
1..1152
/gene="bre-4"

/note="glycosyltransferase; required for susceptibility to
Bacillus thuringiensis Cry toxins"
/codon_start=1
/product="BRE-4"
/protein_id="AA521308.1"
/db_xref="GI:4260122"
/translation="MAFRHLAVARLKSLLVLCVALLVHMYTKIPSLYENLTIGSS
LADVDAMEAVLGNASTSDLDLTWNSTFSPISVENQTSFMEDIRPILFPNQTLOF
CNOEPHVLVGPRIKFLDEPDKLEKIVDTAGGGHMPKDCVARRRVAIVYDRE
ALRILMNLHSLAKQQLDVAIPVEOVANOTFRGKLMNGYVASRLYWCQETP
HDVLLPDDRLATCPRIOPRRMSVAIDENYKLPYSAPGSIATLTKDKLKNGFS
NDFWKGESDDDLATRTSMAGLKVSRYPQIARVYKMTKSTEAIVPNKCRKIMQGT
KRRWTRDGLSNLKYKLVNLELPLVYTRAVVDLEKDCRELRRDPPTCF"

ORIGIN

Alignment Scores:

Pred. No.: 2,84e-179 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-661-430-1 (1-383) x AY533306 (1-1152)

QY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuSerLeuLeuValLeuCysAlaVal 20
DB 1 ATGGCTTTTCCTCATTTGGCAGTCGCCAGACTCACTGCTTCTGCTTCTTGTGCGGTT 60
QY 21 LeuLeuLeuValHisAlaMetIleTyrIleProSerLeuTyrGluAsnLeuThrIle 40
DB 61 CTTCTATTAGTTCATGCAATGATTTTATAGATTCCTTACGAGAACCTTACTATTC 120
QY 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlySerThrAla 60
DB 121 GGCTCCTCCAGCCCTTATTCGCCAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 61 SerThrSerAspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluVal 80
DB 181 TTCACCTCGGATGATCTTATGATACGTGAATTCACGTTTCCAGCAATTCCTGAGATT 240
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
DB 241 AATCAGACTAGTTTATGAGAGACATTCCTCAATCTGTTCCCGACCAACAGACTCTT 300
QY 101 GlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGlu 120
DB 301 CAATTCCTGTAATCAGACACCTCCCACTGTCGAGACCATCGGTATTTCTCGATGAG 360
QY 121 ProAspPheLeuThrLeuGlnIleTyrProAspThrHisAlaGlyGlyHisGlyMet 140
DB 361 CCCGACTTAAACTCTCGAGAAATCTATCCGAGACCCAGCCGCGTGAACATGGAATG 420
QY 141 ProLeuAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGlu 160
DB 421 CCTAAGAGATTGTTGCAAGCATCGTGTCTATATTGTCCTTACCTATGAGATCGTGA 480
QY 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnGlnLeuAsp 180
DB 481 GGACACTTGAAGATATGCTCCACATTTGCACTGCTGCTGCGCAAAACAAATTTGAC 540
QY 181 TyrAlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgIleTyrLeuMet 200
DB 541 TATGCAATTTCTATCTGTGAGCAAGTGGCAATTCAGACTTTATATGCGGAAACTAATG 600
QY 201 AsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAsp 220
DB 601 AACGTTGATACGACATCAACCTCTACCCATGAGAGTGTTCATCTTCTCATGAT 660
QY 221 ValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHis 240
DB 661 GTGCAATTACTCCCGAGAGATGACGTAACCTGTACAGTGTTCATTCACACGTCAT 720

QY 241 MetSerValAlaIleAspLeuPheAsnTyrIleLeuProTyrSerAlaIlePheGlyGly 260
DB 721 ATGAGTGTAGCGATCGATTAATTCATTAATACTTCATATTCGGCGATCTTCGGCGGA 780
QY 261 IleSerAlaLeuThrLysAspHisLeuLysIleAlaGlnIlePheSerAsnAspPheTrp 280
DB 781 ATCACTGACATCAAAAGATCACTGAAGAAATCAATGATTTTCGATATATTTTGG 840
QY 281 GlyTPGlyGlyGluAspAspAspLeuAlaThrArgTrpSerMetAlaGlyLeuLeuVal 300
DB 841 GGTGGGGCGAGAGACGACGATTTGGCGAGCAACATCATGCTGACTGCTGAAGATT 900
QY 301 SerArgTyrProThrGlnIleAlaArgTyrIleMetIleLysHisSerThrGluAlaThr 320
DB 901 TCAAGATATCCGACACCAATTCACAGATTAATAAGATTAGCACTCGACGGAACGACG 960
QY 321 AsnProValAsnLeuCysAspArgTyrIleMetGlyGlnThrLysArgArgTrpThrArg 340
DB 961 AATCCAGTTAATTAATATGCGCTACAAATTAATGGCCCAACGACGCGATGACACGT 1020
QY 341 AspGlyLeuSerAsnLeuLysTyrIleLeuValAsnLeuGluLeuLysProLeuTyrThr 360
DB 1021 GAGGCGCTTAGCAATCTGAATTAAGCTCGTAATCTGAAATTAAGCTCTTCACT 1080
QY 361 ArgAlaValAlaAspLeuLeuGlnLysAspCysArgArgGluLeuArgAspPhePro 380
DB 1081 CGAGCGCTGCTGATTTGCTCGAATAAAGACTGCCCGGAGCTCGAAGGACTTTCCA 1140
QY 381 ThrCysPhe 383
DB 1141 ACGTGTCTT 1149

RESULT 3
AC025727/c 63747 bp DNA linear INV 27-MAY-2005
LOCUS AC025727
DEFINITION Caenorhabditis elegans cosmid Y73E7A, complete sequence.
ACCESSION AC025727
VERSION AC025727.3 GI:14530915
KEYWORDS HTG.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 63747)

REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
PUBMED
9851916
2 (bases 1 to 63747)
AUTHORS
Du,H. and Maupin,R.
TITLE
The sequence of C. elegans cosmid Y73E7A
JOURNAL
3 (bases 1 to 63747)
AUTHORS
Waterston,R.H.
TITLE
Direct Submission
JOURNAL
Submitted (12-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 63747)
REFERENCE
AUTHORS
Waterston,R.
TITLE
Direct Submission
JOURNAL
Submitted (13-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 63747)
REFERENCE
AUTHORS
Waterston,R.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUN-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
6 (bases 1 to 63747)
REFERENCE
AUTHORS
Waterston,R.

TITLE Direct Submission
JOURNAL Submitted (03-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 7 (bases 1 to 63747)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 8 (bases 1 to 63747)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 9 (bases 1 to 63747)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 10 (bases 1 to 63747)
AUTHORS Wormbase Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 11 (bases 1 to 63747)
AUTHORS Wormbase Consortium
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 12 (bases 1 to 63747)
AUTHORS Wormbase Consortium
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
COMMENT On Jun 23, 2001 this sequence version replaced gi:10800387.
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jes@sanger.ac.uk
NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.
This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m3 subclone.
For a graphical representation of this clone sequence and its analysis see:
http://www.wormbase.org/db/seq/sequence?name=Y73E7A;class=Sequence
NEIGHBORING CLONE INFORMATION
The 5' clone is Y92H12A, 200 bp overlap; the 3' clone is Y71G12B, 4000 bp overlap.

NOTES:
Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yui Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (<http://world.bcfi.harvard.edu/>), similarity to other proteins from Blast analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's MABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans Genbank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucleic Acids. Res., 25, 955-964).
Location/Qualifiers
1. 63747
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y73E7A"
5487..12773
/locus_tag="Y73E7A.3"
join(5487..5659,7957..8063,8778..8971,9035..9119,10020..10254,11408..11615,11668..11742,11836..11943,12669..12773)
/locus_tag="Y73E7A.3"
/standard_name="Y73E7A.3"
/note="contains similarity to Pfam domain PF06027 (Bukaryotic protein of unknown function (OUF9141)); coded for by the following C. elegans cDNAs: YK1142a05.3, OSTF30045C02"
/codon_start=1
/product="Hypothetical protein Y73E7A.3"
/protein_id="AAG23391.3"
/db_xref="GI:60543518"
/db_xref="WormBase:Y73E7A.3"
/translation="MTNCGAESTSAPPQANGHLRPPGIQPHRNPPTPPSPDEPDDAINGSPCCDDPVRVSRFKALILGILSLCLCGVSSQLANONVNAAPAOAFSNFLICPVYISLACNDENGLVYVVRKGRWRLLIALLIDVENWYIVAYOYTNLTSVOLDCATPTVLPFLSVRYLASHILGVITCIIGIACVYMAOLDGKMLDGSNNYLGILCAAAMVMAICNVAEFLVKHRSREYVGMGLFECTIVSGVTAVPEOALSKIWDGTVSYFALFAPSMFIFYSILVTVLQKTSALMNFSLTADAFYSLLFGIEMFDOTFHYLYVSRIVLCIGSIYSKKEQVADDEPRRVPCIFVCCCGCFEEGDSYEGSIDVSPSPREROMGLNPRGSSPCPVHGSQQLSHL"
complement(14300..15495)
/locus_tag="Y73E7A.4"
complement(join(14300..14457,15222..15495))
/locus_tag="Y73E7A.4"
/standard_name="Y73E7A.4"
/note="contains similarity to Pfam domain PF05835 (Synaphin protein); coded for by the following C. elegans cDNAs: YK1215a10.5, YK1728h08.3, OSTF30035H07"
/codon_start=1
/product="Hypothetical protein Y73E7A.4"
/protein_id="AAG23387.1"
/db_xref="GI:10800392"
/db_xref="WormBase:Y73E7A.4"
/translation="MAGFLMKQVGNQLSEVYTGGLMKDGGKRTGEDPPEYIARLEOBERKEKRRKREKRRQGIROKVAIKKKEBVANDPFGRIQGPRTPEETAAEMVAEDDSLIGQLTEQVEKAKTAVTAFEYVKGFPFGK"
complement(19944..20816)
/locus_tag="Y73E7A.5"
complement(join(19944..20130,20270..20541,20589..20816))
/locus_tag="Y73E7A.5"
/standard_name="Y73E7A.5"
/note="contains similarity to Pfam domain PF07735 (F-box associated)"
/codon_start=1
/product="Hypothetical protein Y73E7A.5"
/protein_id="AAG23386.1"

```

/db xref="GI:10800391"
/db xref="MormBase:Y73E7A.5"
/translation="MEOH1FIKFOIIGDIFRLSYOSALSHPLFRKCHFLKLGEGSYI
TTEDMKFLIKNPNLKYFLTRBELRKTIDIKLVNIPRLDPKSGILDLDDKMNCE
KMRFDLNTTIKAREINKTIRYLDGNLPLNRRLQNTMHHWAGDLGLPYSKNPK
RRARIFCDLEFVDCSYGRDIERSDGLATLIENFAMVFLVMHDRFPTDPTLPHFS
NSVSVPTNN"
21160..26092
/locus_tag="Y73E7A.2"
/locus(21160..21238,21285..21638,23801..24363,25697..26092)
/locus_tag="Y73E7A.2"
/standard_name="Y73E7A.2"
/note="contains similarity to Rattus norvegicus AATP
protein (Apoptosis antagonizing transcription factor).";
TR:Q9QYW0; coded for by the following C. elegans cDNAs:

```

Alignment Scores:

```

Pred. No.: 9,11e-107 Length: 63747
Score: 1275.00 Matches: 381
Percent Similarity: 22.50% Conservative: 0
Best Local Similarity: 22.50% Mismatches: 2
Query Match: 62.78% Indels: 1312
DB: 2 Gaps: 5

```

US-10-661-430-1 (1-383) x AC025727 (1-63747)

```

OY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuLysSerLeuLeuValLeuCyAlaVal 20
Db 45925 ATGGCTTTTCGATCTTGGCAGTCGCCACACTCAAGTCGTGCTGATCTTGTCGCTT 45866
OY 21 LeuLeuLeuValHisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThrIle 40
Db 45865 CTTCTATTAGTTCATGCATGATTTATATAGATTCATGCTTTACAGAGACCTTACTATC 45806
OY 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGluValThrAla 60
Db 45805 GGCTCCTCCACCTTATGCCCACGTCGACGCAATGAGAGCAGTGTCCGGGAATACGGCT 45746
OY 61 SerThrSerAspAspLeuLeuAspThrTPaanserThrPheSerProIleSerGluVal 80
Db 45745 TCCACTTCGAGATGATCTACTGATACGTGAAATTCACGTTTTCACCGAATTCGAGTT 45686
OY 81 AengInThrSerPheMetGlu----- 87
Db 45685 AATCAGACTAGTTTATGGA -GGTAAGGTTCTTAAACTTATTGGAATTTTGGAA 45627
OY 87 ----- 87
Db 45626 AAAAAATTTTTCGAAATTTTGAATAATTTGAATTTTCTCCAAATCTCAAA 45567
OY 87 ----- 87
Db 45566 AATTCACGTGTCCTGACGACTGTCCTTTACGTTGATCTACAAAAATGCGGAAT 45507
OY 87 ----- 87
Db 45506 TTTTGGCCAAAAATGTGACCTCAGCAGACGTTCTTAACATGCAAAATCAGTGAGAA 45447
OY 87 ----- 87
Db 45446 CTTTCGCGCTCTTTCCTCCGCAATTTTGTAGATCGATGATCAAGCGAAATGGGGG 45387
OY 87 ----- 87
Db 45386 CTTTGAGACCGGCTGTATTGTAAAAAAATTCAGAGAAATTTCAAAAAAATTGGAA 45327
OY 87 ----- 87
Db 45326 GTATATAAAAGTTTTTTTCCAAATTTGAATAATCTGAATATTCCAAATCTTTC 45267
OY 87 ----- 87
Db 45266 AAAAAATTTTCTTTTGAATTTTATTGAAAAAACAATTCGTGTTTATT 45207

```

```

OY 87 ----- 87
Db 45206 TTTTGAATTAATAATAATATATGTCACAAATTTTCCAAATATCCAAAAATTGG 45147
OY 87 ----- 87
Db 45146 AATTTCTCMAAAAAAAGAAATTAATTCMAATTTGTCAGTGAGTACACAATT 45087
OY 87 ----- 87
Db 45086 CGTCTTTCTTTTCAATTTGGCCAGCCGCAAAATGCTAATCAGAAATTTTGGG 45027
OY 87 ----- 87
Db 45026 GAATTTGAAAAATTTGAATTTTAAAAAAATTTTCTGTTTAAAAAAATTAA 44967
OY 87 ----- 87
Db 44966 AATTTCTTTTAAAAAAGCACAACAGACAAATTTTACAGGATCTGACTTCCT 44907
OY 87 ----- 87
Db 44906 CACATGTTGAATGAGAGATTGCGCAACTAGCGCAGCCATATCTGGGTGATTAC 44847
OY 87 ----- 87
Db 44846 GCGCGATTTACGGCGCACTACTTTCCAGTCCGACAGCGATTCATGAGAGCGGA 44787
OY 87 ----- 87
Db 44786 AAATTCATGAGAGAGCCAGAACCCCATGTAATCTAGTTAAATAAAATTTGAAAA 44727
OY 87 ----- 87
Db 44726 AAAAAATTTGAGTTGAGAAAAATTTTGAATAATTCAGAAAAAATTTGTGTAACCTTT 44667
OY 87 ----- 87
Db 44666 TGAATTTTATATTAAGGAAACACAAATTTCTGAGAAATCGCTACTGACACATATTGA 44607
OY 87 ----- 87
Db 44606 CGCGCAATATCTGTAGCAAAAACTGCAATCATTTGACTAGTGGCTGTGTGCAC 44547
OY 87 ----- 87
Db 44546 TTACGGCTCGTTTGATACAGTAGTCATTGAAAAAGAAATTACTGATGTTTCGTAAG 44487
OY 87 ----- 87
Db 44486 ATATTTTCGGGCTCAAAATATGTGTGTATACGCAATTCACAAATTTTCTGTCCATTA 44427
OY 87 ----- 87
Db 44426 TAAAAGTAAAAAATTAATTTATCTTTAGTAAAGGATTTTATAGCGTAAGATT 44367
OY 88 ----- 87
Db 44366 GAGCATTAATAACAATTTTCAAAAAAATCCACAATTTCCAGAAATTCGTAATCCG 44307
OY 94 PheProAspArgInThrLeuGlnPheCysAengInThrProProHisLeuValGlyPro 113
Db 44306 TTCCCGACAAACAGACTCTTCAATCTGTATATCGAACCTCCCACTGTGTGGACCC 44247
OY 114 IleArgValPheLeuAspGluProAspPheIleThrLeuGluLysIleTyrProAspThr 133
Db 44246 ATCCGTGATTCCTCGATGAGACCCGACTTCAAACTCTCGAAGAAATCTATCCGACAG 44187
OY 134 HisAlaGlyValHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 44186 CAGCGCGGTGACATGAAATGCTTAAGGATTTGTGTGAAGGCATCGTGTGCTATTATT 44127

```

QY 154 ValProTyrArgAspArgIuAlaHisIleu----- 163
Db 44126 GTCCTTAAGAGATGCTGAGACATTTGAGGTAGAAATTCGATGCTGGGCTCT 44067
QY 163 ----- 163
Db 44066 GGCACGATTACGGAATTTGAGAAATTTGAAAAAAATTTGAAAAAAATTTT 44007
QY 163 ----- 163
Db 44006 TAAATAAATTTTGAAATTTTAAAAATCAAAATTTCAAACTTGACGATTCAAAAACAT 43947
QY 163 ----- 163
Db 43946 GTTAGGTATCATTTTGGGTCTGGCTGCAATTAACGAATTTTCAGAAATTTGAACCTTT 43887
QY 163 ----- 163
Db 43886 TTGGAAACGAAAAAATTTGTGAAATTTGTAAATTTTTCGATTTTAAAA 43827
QY 163 ----- 163
Db 43826 TTTCAAAATTTCAAAATCCCAAACTGTGATTTTGTCCAAAAAAGAGTAATCCA 43767
QY 163 ----- 163
Db 43766 CAATTCGATTCGCGCGAATCGACTTCCCTATAAGTGAATTTTGTTCCGGGGATT 43707
QY 163 ----- 163
Db 43706 CCAACTACGCGACTTTTCCATGAAATTTGGTTTGTAGACATTTTGTAGATAAAT 43647
QY 163 ----- 163
Db 43646 TGAAGCTCAGAAACGAGATTAAGTAAGAAATTTAGAAATTTAGCTCTGCGACGATTT 43587
QY 163 ----- 163
Db 43586 ATAAAGATTGAAAAAATTTGAAAAAGAAATTTTGTGAAAAATTTGAAAAAAGTTAA 43527
QY 163 ----- 163
Db 43526 AAAAAATTTTGTGTCATTAATTTCCATATTTATCATAGTAATATATACCTTTTC 43467
QY 163 ----- 163
Db 43466 AAAAAATTCCTCACTCACTGCTGTCAAGTGTCTCATTTGTGATCTACGTTGATCTA 43407
QY 163 ----- 163
Db 43406 CAAAAATGCGGGAAGAGGGGCAAGAGTTCTCAAGATTTGCGCATGTTAAGAACGTG 43347
QY 163 ----- 163
Db 43346 CTGACGTCAATTTTGTGGCAAAAAATTCGCGCATTTTGTGATCAAAATCTGAT 43287
QY 163 ----- 163
Db 43286 GAGACAGCTGGACCACTGTGCTCAATGAAATTCGAAACCCGTGATTTCAACTACCCC 43227
QY 164 ----- 174
Db 43226 GATTTTCCGATTTGAGCTATATTTTTCAGAAATATGCTCCGCAATTTGCACTGCTCTC 43167
QY 175 AlAlaYgInglInleuAspTyrAlaIlePheIleValGluInValAlaAsnGlnThrPhe 194
Db 43166 GCCAAACAAACATTTGACTATGATTTTCATGTGAGCAAGTGGGAAATCAAGACGTTT 43107
QY 195 AsnArgGlyIys----- 198
Db 43106 AATCGCGGAAAGTATGTTATTTATTTATTTTCTCCAAATTTCTAAACAAATTTGGA 43047
QY 198 ----- 198

Db 43046 GAAAAAATGATATATATCTTTGGGGTCCCGGTGACGTCACGCGTGGCCGTGACTAAAAAA 42987
QY 198 ----- 198
Db 42986 ATCAAAAAATAGGTAGGGGACCCACGTGTGTACAGCTGTCCAAATACAGTCTGATCT 42927
QY 198 ----- 198
Db 42926 AAAAAAATGCGGAATTTTTCCAAATTTGTAGCTATCAAGCTCTTAATTATGCAAA 42867
QY 198 ----- 198
Db 42866 ATCAGTTGAAACTGTGCTCACTATTTTTCGAAGATCAAGCGAATGGAACCTTT 42807
QY 198 ----- 198
Db 42806 TGACTCCAGTGGAGCCCTCTGACGTACGCGTGGCCGTGACTAAAAATTCAAAAAT 42747
QY 198 ----- 198
Db 42746 AGAGTTGGGGACCCATCTGACGTCAATGTGACATCATAGAGGCGCCGCGGCTTAGA 42687
QY 199 ----- 209
Db 42686 ATCTCCAAATATAAAAATTTCTTCAGCTAATGAACGTTGATACGACGTAGCATCACGC 42627
QY 210 LeuTyrProTyrGlnCysPheIlePheHisAspValAspLeuProGluAspAspArg 229
Db 42626 CTCTACCAATGCGAGTGTCAATCTTTCATGATGATGATTTACTGCCCCAAGATGACCGT 42567
QY 230 AsnLeuTyrThrCysProIleGlnProArgHisIleMetSerValAlaIleAspYspPheAsn 249
Db 42566 AACCTTACACGTGTCCATTTCAACCACTCATATAGTGTACGATTCATATAATTCAT 42507
QY 250 TyrIysLeu----- 252
Db 42506 TATTAAGTA-GGTTTTTCTGCAATAAAAAATTTCAAAAAATCTGAAAAAATCGATTCC 42448
QY 252 ----- 252
Db 42447 AACTTTCAAGTCTGCGCTTCTCATGAAATTTTTCGCGTCCATTGACAAATCCCGC 42388
QY 252 ----- 252
Db 42387 CAGGACACGCGGTAGTGTGTACTCCACAGGACAAATACATTAGTTTGCAAC 42328
QY 252 ----- 252
Db 42327 TAAATCGAGCGCGACGCGACAGCAAGCGCGTAAATCTACGTAGCCAAAAATAGCCT 42268
QY 252 ----- 252
Db 42267 AGTTGCGCAAACTCTTCATTTTAATTAATGAGGAAGCAATGATTGAATAATG 42208
QY 252 ----- 252
Db 42207 AAGATTTTGTGAAGTACGCAATTTTGGCTCGGCAATATCTGATGATTTACGGCG 42148
QY 252 ----- 252
Db 42147 CGTTGCGTCCGCGTCCGCTTGAATTTTACTTTGTAACCTAATGATTTGTCCGTGTG 42088
QY 252 ----- 252
Db 42087 GAGTACACGACATTTTCCACGCGTTGTCCGCGGCGCATTTGCAATGACGCGAAAAA 42028
QY 252 ----- 252
Db 42027 TTCAATAGAGAGCGCGAACCCTGTTTAATACATTTTTCGAAAAAATACACAAA 41968
QY 253 ----- 262
ProTyrSerAlaIlePheGlyIleSer 262


```

Db      370 CCATTGTTAATAACCAAGCTAATGATAGCATTAAGATTATGTCGACGATGAAGATGCT 429
Qy      98 -----GlnThrLeuGlnPheCysasnGlnInhrProPheHisleu 110
Db      430 GTGAGAAAGCTGAAGCTGTGTGACGCGCTGCTTGATGATGATGATGATGATGATGAT 489
Qy      111 ValGlyProIleArgValPheLeuAspGluProAspPheLeuGluLeuIleTyr 130
Db      490 ---GGCCCTAATACACTGAACAAACGAACTGGAATGGGGTGAAGAAAAGTTT 546
Qy      131 ProAspThrHisAlaGlyGlyHisGlyMetProLeuAspCysValAlaArgHisAlaGlyVal 150
Db      547 CCGAGAGGTGAGTGGGAGGCGCGTACTCCGCCCCCACTGCAACGCGCCCACTGAAGATC 606
Qy      151 AlaIleIleValProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnLeu 170
Db      607 GCCATTATGTCGATACAGAGATCGCACAGCATTTAGCAATATTCCTGAAATCATATG 666
Qy      171 HisSerLeuLeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGlnGlnValAla 190
Db      667 CATCCGTTTTTGTATGAAGAACGCAAAATTGATATGGAATTTTCAATCGTCGAGCAAGAAAGA 726
Qy      191 AsnGlnThrPheAsnArgIleGlyLeuMetAsnValGlyTyrAspValAlaAsnArgLeu 210
Db      727 AACAGAGACTTCAACCCGCAAGCTGATGAAGCTTGGCTTGGAGAGCCCAAGAGCTG 786
Qy      211 TyrPro-----TyrGlnCysPheIlePheHisAspValAspLeuLeuProGluAspAsp 228
Db      787 GTGGCGAAGAGCTGGACAGTCTTCTGTTCCAGACATTCGATCTTGCTCTCGACACC 846
Qy      229 ArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPhe 248
Db      847 AGGAACTGTAATCTCTGCGCCACGACGCGGACATATGTCGGCTTCATCGATTAATTA 906
Qy      249 AsnTyrLeuLeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHis 268
Db      907 CATTTCAAGTTACCATTCAGAGATATTTTGGCGGAGTCTCTGCGCAAGACCTTGGAAACAG 966
Qy      269 LeuLeuLysIleLeuGlnGlyPheSerAspAspPheThrGlyTyrThrGlyGlyIleAspAsp 288
Db      967 TTCACAGAGGTCACAGGCTTCTGCAATAGTACTGGGGTTGGGGCGGGGAAGATGATGAC 1026
Qy      289 IeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAla 308
Db      1027 ATGTCATATAGGTTAAAGAAATTAATATCATATACATACAGATACCAAAATGTCTATGCG 1086
Qy      309 ArgTyrLysMetIleLysHisSerThrGlnAlaThrAsnProValAsnLysCysArgTyr 328
Db      1087 CCGTACCCCTATGTTAGTATCAAGAAATCTAACACCTATCTTAAG-----AGGTAC 1137
Qy      329 LysIleMetGlyGlnThrLysArgTyrThrArgAspGlyLeuSerAsnLeuLysTyr 348
Db      1138 CAACCTTATACGCAACCAAGCAAACTTCAGAAAGATGCTCTGTGACCGCTCGAGTAT 1197
Qy      349 LysLeuValAsnLeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeuGln 368
Db      1198 GAGTTATCTCAGAGGTGTTCACTAGTACCACTTACACCAATACTAGTGAACATTGACAGAG 1257
Qy      369 Lys 369
Db      1258 CGT 1260

```

```

RESULT 5
BC077601 2265 bp mRNA linear VRT 12-OCT-2004
DEFINITION Xenopus laevis UDP-Gal:betaGalCMC beta 1,4-galactosyltransferase,
polypeptide 3, mRNA (CDNA clone MGC:84507 IMAGE:5077911), complete
cds.
ACCESSION BC077601 GI:50417572
VERSION BC077601.1
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis

```

```

REFERENCE
AUTHORS
1 (bases 1 to 2265)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
Dev. Dyn. 225 (4), 384-391 (2002)
12454917
2 (bases 1 to 2265)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Usdin,T.B., Toshiyuki,S.,
Scheetz,T.E., Brownstein,M.J., Umeda,T.B., Loquellano,N.A., Peters,G.J.,
Carinuci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., C., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Sklisa,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 2265)
Klein,S. and Gerhardt,D.S.
Direct Submission
Submitted (19-JUL-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgsabbs-remail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lise Prabhu,
Parvaneh Saeedi, JR Santos, Angelica Schnerch, Ureila Skalske,
Darane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAX Plate: 152 Row: n Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
source
location/Qualifiers
1..2265
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:84507 IMAGE:5077911"
/tissue_type="Oocytes"
/clone_lib="NICHD_XGC_001"
/lab_host="DH10B"

```



```

gene
    /note="Vector: PCMV-SPORT6.1"
    1..2265
    /gene="D4gal13-prov"
    /note="synonym: MGC84507"
    /db_xref="GeneID:446718"
    178..1188
    /gene="D4gal13-prov"
    /codon_start=1
    /product="D4gal13-prov protein"
    /protein_id="AAH77601.1"
    /db_xref="GI:50417573"
    /db_xref="GeneID:446718"
    /translation="MALYREGYRKRVAFFGIPYKGSASNGILNPHNLGQADVVA
    NLSLSKPSLREELPYCEPTSPFGAPVLPENLLLEVERKNPVYTKGARYKP
    DCESTHKTAVIIPHNREOHLYLLYHAPLOROLYGIYIIHQAGNPFNRKXKL
    NVGFKAEMKDEMDCLFPHVDLIPEDRNITICRPFKHSIANDKGYLPLYSYF
    GVSALSPQYKMGKGFPMNTMGWGEEDDIGIRVALSGMLISRSVQYGRKMTKHG
    HDKNEONPKRFNMLTKTRRTWKODGMNALOYLILSKELQPLTYNIITVDIKGLRP
    HT"

ORIGIN

Alignment Scores:
Pred. No.:      7.56e-59      Length:      2265
Score:          741.00       Matches:      145
Percent Similarity: 59.95%    Conservative: 78
Best Local Similarity: 38.98%  Mismatches:  115
Query Match:     36.48%      Indels:       34
DB:              5           Gaps:         6

US-10-661-430-1 (1-383) x BC077601 (1-2265)

QY 4 ArgH1sleuAlaValAlaArgLeuLysSerLeuLeuValLeuCyAlaValLeuLeu 23
DB 115 AAAAAAATGCGATCTCTTAAGGCTCGCTACTCTTCTTCTCTGCGTCCAGCTAAG 174
QY 24 ValHisAlaMetCileTy-----LysIleProSerLeuTyArgLys 37
DB 175 GTCATGGCTCTCTTATATGAGAAAGCTACAGAAAGGCTGCTTATTTCTTGAGATA 234
QY 38 LeuThrIleGlySerSerThrLeuIleAlaAspValAlaMetCilAlaValLeuGly 57
DB 235 TTTTACAAAGGGGCGCTGCTCT-----TCCATGGAGCTACCAACCC 276
QY 58 AenThrAlaSerThrSerAspAspLeuLeuAepThrTriPAsSerThrPheSerProile 77
DB 277 CACAAATTATCCCAAGCTGGGACGTT-----TATCCCAACCTC 315
QY 78 SerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsn 97
DB 316 AGCTTAATATCCAAAGCTTCGCTGGGGAA----- 345
QY 98 GlnThrLeuGlnPheCyAsnGlnThrProProHisLeuValGlyProIleArgValPhe 117
DB 346 GAGGAGTTGCCCTACTGCTCTAGACCTCTCTTCTTGAAGGGGCGCAATAAGAGTCAG 405
QY 118 LeuAspGluProAspPheLysThrLeuGluLysIle-----TyrProAspThrHis 134
DB 406 TTG-----CCGAGAACTTGTACTGAGAGAGTCGAACGAAAAATCTTAACTGACC 459
QY 135 AlaGlyGlyHisGlyMetCProLysAspCyValAlaArgHisArgValAlaIleIleVal 154
DB 460 AAGAGTGGCGGTACAAAGCGCCGACCTCGAGTCACCAACCAATAAATCGCGTATCANTC 519
QY 155 ProTyraArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeu 174
DB 520 CCTCACCGGACCGGACGACGACCTTAATATCTGCTCTACTACTTGCAACCGCTTTCG 579
QY 175 AlaLysGlnGlnLeuAspTyraIaIlePheIleValGluGlnValAlaAsnGlnThrPhe 194
DB 580 CAGCGGACCAACTCAACTACCGGATATATCATTCATCAGCAGCATATTCACCTTTC 639
QY 195 AanaArgGlyLysLeuMetAsnValGlyTyraAspValAlaSerArgLeuTyraProTriP 214

```

```

DB 640 AACGAGCAAGATTACTGAACTCGGCTTCAAGAGGCCATGAAGACGAGACTGGAC 699
QY 215 CysPheIlePheHisAspValAlaAspLeuProGluAspAspArgAsnLeuTyThrCys 234
DB 700 TGCCGTCTTCTTCCACGATGCGACCTCTATCCCGAAGCGATCGTAACATTTATATAGT 759
QY 235 ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyLysLeuProTy 254
DB 760 GACAGATTTCCAAAGACAGCGCTCCATGCGCAGTGAACAAGTTTGATACAAGTTCCTTAC 819
QY 255 SerAlaIlePheGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGly 274
DB 820 AAATCTCTCTCGGAGGAGTGTCTGCGCTGAGCCGACGCAATCATCAAGATGAATGA 879
QY 275 PheSerAsnaAspPheTrgLIITrGlyGlyLysAspAspLeuAlaThrArgThrSer 294
DB 880 TTTCCAAACAACCTACTGGGGCGGGGGGTGAAGTGCACATATGGCATCAGAGTGGCA 939
QY 295 MetAlaGlyLeuLysValSerArgTyraProThrGlnIleAlaArgTyLysMetCileLys 314
DB 940 CTCAGTGCATGCTTATATCTCGCCCTCAGTCCAGTACGGGAGATATAAATGATCAAG 999
QY 315 HisSerThrGluAlaThrAsnProValAsnLysCysArgTyTyLysIleMetCileGln 334
DB 1000 CACGCGCATGACAAAGCGCAATGACGAAATCCCAAAAGTTCAACATGTGACTAAGACT 1059
QY 335 LysArgArgTyraThrArgAspGlyLeuSerAsnLeuLysTyTyLysLeuValAsnLeuGly 354
DB 1060 CGAGAGACATGAGAACGATGAGATGAACGACATGCTATCATATCTGATCTCAAGAG 1119
QY 355 LeuLysProLeuTyTyThrArgAlaValAlaAspLeu 366
DB 1120 CTGACGACCATATACCAATATAACGTGACATT 1155

RESULT 6
CR942597      2098 bp      mRNA      linear      VRT 18-APR-2005
LOCUS        Xenopus tropicalis finished cDNA, clone TGa8115g12.
DEFINITION   CR942597
ACCESSION   CR942597.1 GI:62732412
VERSION      CR942597.1 GI:62732412
KEYWORDS
SOURCE
ORGANISM     Xenopus tropicalis (Silurana tropicalis)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 2098)
Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
Voigt,J. and Zorn,A.M.
Direct Submission
Submitted (15-APR-2005) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST/cDNA project.
This sequence is from a Xenopus Gene Collection (XGC) library, from
a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.

FEATURES
    source
        1..2098
        /organism="Xenopus tropicalis"
        /mol_type="mRNA"
        /db_xref="taxon:8364"
        /clone="TGa8115g12"
        /clone_lib="XGC-gastrula"
        /dev_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN

Alignment Scores:

```


Pred. No.: 8,53e-59 Length: 2098
 Score: 740.00 Matches: 145
 Percent Similarity: 60.22% Conservative: 76
 Best Local Similarity: 39.51% Mismatches: 118
 Query Match: 36.44% Indels: 28
 DB: 5 Gaps: 5

US-10-661-430-1 (1-383) x CR942597 (1-2098)

```

QY 6 LeuAlaValAlaArgLeuYsSerLeuValLeuCySaValLeuLeuValHis 25
    ::::::::::::::::::::
DB 13 ATGCTCATGCGCTAGCGCTACCTTCTCTCTGCGCAGCTGGTGCTATG 72
QY 26 AlAmetIleTyrlsIleProSerLeuTyrluAsnLeuThrIleGlySerSerThrLeu 45
    ::::::::::::::::::::
DB 73 GCTCTGCTGTATAG-----GAGGCTACAGAAAAG 105
QY 46 IleAlaSprValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAspAsp 65
    ::::::::::::::::::::
DB 106 GTGGCT-----TATTTCTTGGTATATTTTACAAGGGGACCTGCTCCATGGA 156
QY 66 LeuLeuAspThrTrpAsn-----SerThrSerSerProIleSerGluVal 80
    ::::::::::::::::::::
DB 157 CTGATGAAACCCACATGTATCCAGCCCGGACGCTTACGCCAACCTCACCTTATA 216
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
    ::::::::::::::::::::
DB 217 TCCAGCTCTCTTACGGAA-----GAGGACTTG 246
QY 101 GlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGlu 120
    ::::::::::::::::::::
DB 247 CCATTCCTGGCGGAGCTCACTTATATAGGGGCGCCAAATAGAGTCAGCTTCCGAG 306
QY 121 Pro---AspPheLeuThrLeuGluValIleTyrlProAspThrHisAlaGlyHisGly 139
    ::::::::::::::::::::
DB 307 AACTTGTACTGGAGGAGGTGAGAAAAGAACCTTACCTGTCGCAAGAGGCGCTTAC 366
QY 140 MetProLysAspCysValAlaArgHisArgValAlaIleIleValProTyrlArgAspArg 159
    ::::::::::::::::::::
DB 367 AAGCGACCCGACCTCGAGTCCACATCAAAAGCGCGCTGATCATCCCTCACCGCGAGG 426
QY 160 GlnAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnGlnLeu 179
    ::::::::::::::::::::
DB 427 GAGCAGACCTTAATATACCTGCTATATATCTGACCCGTTTCTGCGACGCGCAACTG 486
QY 180 AspTyrlTrpAlaIlePheIleValGlnGluValAlaAsnGlnThrPheAsnArgGlyLeu 199
    ::::::::::::::::::::
DB 487 AACTACGGGATATCATCATCATCAGCGGGTAAATTTTACTTTCACCGCGCTAAGGTG 546
QY 200 MetAsnValGlyTyrlAspValAlaSerArgLeuTyrlProTyrlGlnCysPheIlePheHis 219
    ::::::::::::::::::::
DB 547 CTGAACTCGGCTTCAAGAGAGCCATGAAGAAGATTGGGACTGCTGTTTCTTACAC 606
QY 220 AspValAspLeuLeuProGluAspAspArgAsnLeuTyrlCysProIleGlnProArg 239
    ::::::::::::::::::::
DB 607 GATGTCAACCTCATTTCCCGAAGAGATCGCAATATCTTACCTCGCAAAATTTCCCAAG 666
QY 240 HisMetSerValAlaIleAspIlePheAsnTyrlsLeuProTyrlSerAlaIlePheGly 259
    ::::::::::::::::::::
DB 667 CACCCCTCGATCGCATGAGCAAAATTGGATACAAATTTGCTTAAATCTTACTTGCGG 726
QY 260 GlyIleSerAlaLeuThrIleAspHisIleLeuTyrlsIleAsnGlyPheSerAspAspPhe 279
    ::::::::::::::::::::
DB 727 GGAATGTCAAGCTTTTCCCGAAGACATATCATGAAGTGAAGCGCTTCCCAACAATCAC 786
QY 280 TrpGlyTyrlGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLys 299
    ::::::::::::::::::::
DB 787 TGGGAGATGGGGCGGAGGAATGATGATATAGCATCAGGGTCGCACTTACGCGCATGATC 846
QY 300 ValSerArgTyrlProThrGlnIleAlaArgTyrlsMetIleLysHisSerThrGluAla 319
    ::::::::::::::::::::
DB 847 ATCTCTGCGCTTCTATTCAGACAGCGGAGGTATAAATGATCAACGAGCGCATGACAAAG 906
  
```

```

QY 320 ThrAsnProValAsnLysCysArgTyrlsIleMetGlyGlnThrIysArgTrpThr 339
    ::::::::::::::::::::
DB 907 GCGAATGACCAAAATCCCAAAAGTTTAACTATACAGAACAGCCGCGGAGCGG 966
QY 340 ArgAspGlyLeuSerAsnLeuTyrlsTyrlsLeuValAsnLeuGlnLeuLysProLeuTyrl 359
    ::::::::::::::::::::
DB 967 CAGGACGGATGAAATCGCTGCAAGTACTCTTCCAAAGAGATTGACGCGCTTAC 1026
QY 360 ThrArgAlaValAlaAspLeu 366
    ::::::::::::::::::::
DB 1027 ACCAATATCATCTGTAACATC 1047
QY 360 ThrArgAlaValAlaAspLeu 366
    ::::::::::::::::::::
DB 1027 ACCAATATCATCTGTAACATC 1047

RESULT 7
BC068719
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 2681)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
Dev. Dyn. 225 (4), 384-391 (2002)
2 (bases 1 to 2681)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,D., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Teshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Adamsion,R.D., Mulhany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Hellon,E., Kerteman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whiting,M., Madan,A., Touchman,J.W., Green,E.D.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 2681)
Klein,S. and Strausberg,R.
Direct Submission
Submitted (02-APR-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
  
```

Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabu, Parranah Seedi, JR Santos, Angelique Scherch, Ursula Skalska, Duane Smallus, Jeff Stott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIN: at: <http://image.llnl.gov>
Series: IRAX Plate: 156 Row: 1 Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source
1. 2681
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:81163 IMAGE:6318052"
/issue_type="Ovary, Xenopus"
/clone_1ib="NICHD_XGC_Ov1"
/lab_host="PH10B"
/note="Vector: PCMV-SPORT6"
1. 2681
/gene="MGC81163"
/db_xref="GeneID:414540"
77. 1573
/gene="MGC81163"
/codon_start=1
/product="MGC81163 protein"
/protein_id="AAH68719.1"
/db_xref="GI:46250086"
/db_xref="GeneID:414540"
/translation="MMKRYLERPCTLAMIIGFQAPMYFRSGRLSLFIRGSHIEP
SDYRSTHDVSNLSVNSWVLPSTAEILIPCEPSYLVGPIVTSQPSLKK
VODKPLVKGGEYKPCNEAHRRTAVIIPIHNRBTIHLIYLAHPJLQRCQALYRI
YIHOAGSTFNRAKLIVGVEALRDEMDCLFLADVDLIPEBNFNLVCCPMSPKH
ASVANKFSEYLPYPMYFGVSGALTPDQKNGKGFENETMGWGSDDDIATVRLAGM
KTRPSVAGHKMKHKGHDHNEBPHFDLIRPMMKTKDGKNSLNTYKLAIELE
HLTYNTVDIGIDPRAKRLRNRGPRPPPIRYNKTSTSDQIRNADQCLTVSE
INGVPSAEKRPVAPBSTTEKSLPLTASAPREKVEDKQIRVIRSTEDNNRAOV
KEVQADVARGRAQRSSKLERQVDSFRAPQDDVKSJLPQDLSRAPV"

ORIGIN

Alignment Scores:

Pred. No.: 2,44e-58 Length: 2681
Score: 736.50 Matches: 149
Percent Similarity: 60.78% Conservative: 54
Best Local Similarity: 44.61% Mismatches: 114
Query Match: 36.26% Indels: 17
DB: 5 Gaps: 6

US-10-661-430-1 (1-383) x BC068719 (1-2681)

QY 55 ValLeuGLyAaThrAlaSerThrSerAaPLeuLeuAaSPThrTrpAaSerThrPhe 74
DB 185 ATCTTCGGAGACTCCTCATTGAGCCAGTTTGACTATTCGACAC---CATATGCTTAC 241
QY 75 SerProILeSerGluValAaGln-----ThrSerPheMetGluAaPLeuArg 90
DB 242 ACCAACCTCTCCCTGCTCAACTCCATGCTGCTCCCTACTAGACACAGCAAGACATTC--- 298
QY 91 ProILeLeuPheProAaPasnGlnThrLeuGlnPheCysAaGlnThrProProHisLeu 110
DB 299 -----CTGCTTCTCTAGATGCCAGAGAGGCTGCACCTTATTTG 334
QY 111 ValGlyProILeAaArgVal---PheLeuAaPLeuProAaPheLeuThrLeuGluValLeu 129
DB 335 GTGGGGCCCATTAACGCTTCAAGCCAAACCCCTCCCTTAAAGAGTTCAAGTAA 394
QY 130 TrpProAaPThrHisAaGlyGlyHisGlyMetProLyAaSPCyValAAlaAaGHisArg 149
DB 395 AACCCCTGTGATATAAAGTGGGAGTACAAGCCCGAAGCTGTGAGGCGCGGACCCG 454

QY 150 ValAlaIleIleValProTrpArgAaPArgLualAaHisLeuArgILeMetLeuHisAa 169
DB 455 ACGGCGGCATCATCTCCTCAGCGAACCAGAGACTCCTGAGCACCCTGCTTATTAAC 514
QY 170 LeuHisSerLeuLeuAlaLyGlnGlnLeuAaPTrpAlaIlePheILeValGluGlnVal 189
DB 515 CTGACACCCCTTCTGCTGACCGCAGACGCTCCACTACAGATCTACATCATCATCAGGCG 574
QY 190 AlaAaGlnThrPheAaAaArgGlyLyLeuMetAaValGlyTrpAaPValAAlaSerArg 209
DB 575 GGTAACTCCACATTTAAATCGGCGCCAAAGCTGGAATGTGGGGTGAAGAAAGCTCGGCG 634
QY 210 LeuTrpProTPGInCysPheIlePheHisAaPValAaPLeuLeuProGluAaPAspArg 229
DB 635 GACGAGACTGGAGACTGTCTGTTTCTCATATGATGATGATCTGAGATATACCTTC 694
QY 230 AaLeuTrpTrpCys---ProILeGlnProAaGHisMetSerValAlaIleAaPLeuPhe 248
DB 695 AACTGTATGTGTGTGACCCCTGAGGCCCAACAGCGCTGTGATCATGAACAAATTC 754
QY 249 AaTrpTrpLeuProTrpSerAlaIlePheGlyGlyIleSerAlaLeuThrLyAaPHis 268
DB 755 AGTTACAACTGCGCATACCTATGATATTTGGGGAGTCTGCTTGAACCCAGACAG 814
QY 269 LeuLySlyIleAaGlyPheSerAaAaPheTrpGlyTrpGlyGlyLyAaPAspAsp 288
DB 815 TATATGAAGATATATGATTTCTTATATGATTTGGGATGGGAGGTGATATATGAT 874
QY 289 LeuAlaTrpArgTrpSerMetAaGlyLeuLySlyValSerArgTrpProThrGlnIleAla 308
DB 875 ATGCGACACAGGCTCGTGGCGCGTATGAATAACACCCGCGCTGCTGCGCTGGCG 934
QY 309 ArgTrpLyMetIleLyHisSerThrGluAlaTrpAaPProValAaLyAaPLeuArgTrp 328
DB 935 CACTTAAAGATGCTCAACACAAAGGAGACCATGCGAACCAATCCACACAGGTTT 994
QY 329 LySILeMetGlyGlnThrLyAaArgTrpTrpTrpAaPArgLyLeuSerAaLeuLySlyTrp 348
DB 995 GACCTTCTTATCAAGAACCCAGAGATGTGAAAGACAGCGGATATACCTCGTGAACATAC 1054
QY 349 LySLeuValAaLeuGluLeuLyPheProLeuTrpTrpArgAlaValAaPLeu---Leu 367
DB 1055 AAGCTGCTGGCCAGAGAGCTGGAACATCTTACACCAACGAGCTGTGATATAGGGG 1114
QY 368 GluLyAaPLeuCyAaArgGluLeuArgAaPheProTrp 381
DB 1115 GACCCGAGGCAAGAAACCCCTCAGAAATCGGGGTCCAAACC 1156
RESULT 8
AK116794
LOCUS AK116794 2082 bp mRNA linear INV 30-NOV-2002
DEFINITION Clona intestinalis cDNA, clone:c1eg035d14, full insert sequence.
ACCESSION AK116794
VERSION AK116794.1 GI:23589473
KEYWORDS FLI CDNA.
SOURCE Clona intestinalis
ORGANISM Clona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Clona.
REFERENCE
AUTHORS
1 Saeou Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T.,
Sasaki, A., Hamaguchi, M., Awazu, S., Yagi, K., Sasaki, Y.,
Nakayama, A., Ishikawa, H., Inaba, K. and Satch, N.
A cDNA resource from the basal chordate Clona intestinalis
Genes 33 (4), 153-154 (2002)
2 (bases 1 to 2082)
Saeou Y. and Satch, N.
Direct Submission
Submitted (04-OCT-2002) Nori Satch, Kyoto University, Department of
Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail: satch@acidian.zool.kyoto-u.ac.jp, Tel: 81-75-753-4095,

COMMENT Fax:81-75-705-1113)
 Clona intestinalis cDNA Project (URL:
 http://ghost.zool.kyoto-u.ac.jp/index1.html).
 FEATURES Location/Qualifiers

source
 1.2082
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="c1eg035d14"

ORIGIN

Alignment Scores:

Pred. No.: 5.22e-58 Length: 2082
 Score: 731.50 Matches: 148
 Percent Similarity: 60.18% Conservative: 53
 Best Local Similarity: 44.31% Mismatches: 120
 Query Match: 36.02% Gaps: 13
 DB: 2 Indels: 4

US-10-661-430-1 (1-383) x AK116794 (1-2082)

QY 36 GUAAsnLeuThrIleGlySerSerThrLeuIleAlaAspValAlaAspAlaMetGluAlaVal 55
 DB 585 GAAATGACAACAATACCTCTACACA-----ACATCTACAGTT 623
 QY 56 LeuGlyAsnThrIleSerThrSerAspAspLeuAspThrTrpAsnSerThrPheSer 75
 DB 624 TTAGGTTCAACA-----ACAACAAGACAACAACAACAACAACAATATCAACAACA 677
 QY 76 ProIleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPhePro 95
 DB 678 ACAACAACAAGACAACAACAACAAGCTGCTCATGAT-----AAACCAACAACAAGCTGA 731
 QY 96 AspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArg 115
 DB 732 GAAGAAGCTATTTTGGCCAGATGTGCTGGAACCCCTCAACCTTACAGGGAATATAGAT 791
 QY 116 ValPheLeu-----AspGluProAspPheLeuThrLeuGluValLeuTyProAspThr 133
 DB 792 GTGCTCTTTGCCACACATGCAACGAGAAATATCCAACTAGAGCGGATACCACTGCTT 851
 QY 134 HisAlaGlyGlyIleGlyMetProLeuAspCysValAlaArgHisArgValAlaIle 153
 DB 852 AAAGATGGCGGTCACTCAAGCCCTTACGTTAAACCTTATGATTAAGTGGCATAGTT 911
 QY 154 ValProThrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
 DB 912 ATTCTTATCGGAGACAGAGAACAATTTGGGATTTTCTCGAGTACATGATCCACC 971
 QY 174 LeuAlaIleGlnGlnLeuAspTyAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
 DB 972 CTCAACAGCTCAACAACCTTGATTAAGCATATGTTATATCAAGCTGAGTAA 1031
 QY 194 PheAsnArgGlyLeuLeuMetAsnValGlyTyArgValAlaSerArgLeuTyProTyr 213
 DB 1032 TTTTAATGTGTCAGACGTAGTAAAGTGTGTTATCTGCAACAAACAAGATCATGATT 1091
 QY 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspAspAsnLeuTyThr 233
 DB 1092 CAATGTTTGTCTTCCATGATGTGACTTGCTGCTTGAAGATGCAAGAGATATATCTCT 1151
 QY 234 CysProIleGlnProArgHisMetSerValAlaIleAspTyPheAsnTyLeuLeuPro 253
 DB 1152 TGCCCTTCAGAGTCTCGGCACCTGTCAAGCTGGGGTGACAAAGTTAATTACCACTGCC 1211
 QY 254 TyrSerAlaIlePheGlyGlyIleSerAlaLeuThrTyAspHisLeuLeuValAsn 273
 DB 1212 TACTCTCCATATTTGGTGGGTGAGCTGACAAAGGAACATTTCCAAAAGGTGAAC 1271
 QY 274 GlnPheSerAsnAspPheThrGlyTyTrpGlyValGluAspAspAspLeuAlaThrArgThr 293
 DB 1272 GGTTACTCTAATTCATTCATTCGGGGTTGGGGTGGAAGAAGCAAGATATGTTTAACCGTGC 1331

QY 294 SerMetAlaGlyLeuLeuValSerArgTyTrpProThrGlnIleAlaArgTyIleMetIle 313
 DB 1332 AAGTTTCTCGGATGAATATATTCCTTACCCTATGATATCTCAAGGATATAAGATGATT 1391
 QY 314 TyHisSerThrGluAlaThrAsnProValAsnLeuValCysArgTyIleIleMetGlyGln 333
 DB 1392 ACTCATCAAAAGAGAGAGAAAGAAACGAACCAACCCAAACGGCTTCCACCAAAATACCGAG 1451
 QY 334 ThrTyArgArgTyTrpThrArgAspGlyLeuSerAsnLeuTyTrpTyHisLeuValAsnLeu 353
 DB 1452 ACTTAAGACACGATGGCAACAGATGCTTAAACACACTTGATACAGAGTTCGATCAAG 1511
 QY 354 GluLeuLeuProLeuTyTrpThrArgAlaValAlaAspLeuLeu 367
 DB 1512 CAGAAACAACAGCTTATATCAATGTAACGGTTGACATTATG 1553

RESULT 9

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

BC041742 2626 bp mRNA linear VRT 04-NOV-2003
 Xenopus laevis similar to UDP-gal-4betaglcNAc beta
 1,4-galactosyltransferase, polypeptide 3, mRNA (CDNA clone
 MGC:52827 IMAGE:4884596), complete cde.
 BC041742
 BC041742.1 GI:27469876
 MGC.

Xenopus laevis (African clawed frog)
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 2626)
 Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
 and Richardson,P.
 Genetic and genomic tools for Xenopus research: The NIH Xenopus
 Initiative
 Dev. Dyn. 225 (4), 384-391 (2002)
 12454917

2 (bases 1 to 2626)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Scheuen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buerow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullany,S.J., Bosak,S.A., McKwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Vallejo,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bonfield,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
 Scherf,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

3 (bases 1 to 2626)
 Klein,S. and Strausberg,R.
 Direct Submission
 Submitted (23-DEC-2002) National Institutes of Health, Xenopus Gene
 Collection (XGC), National Institute of Child Health and Human
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
 20892-7510, USA
 NIH-MGC Project
 NIH-MGC Project
 Contact: XGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

REMARK
 COMMENT

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Jessica Fahay, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLN at: http://image.llnl.gov
Series: IRAX Plate: 94 Row: n Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1..2626

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="MGC:52827 IMAGE:4884596"

/issue_type="Embryo, stage 17/19, Xenopus"

/clone_lib="NICHD_XGC_Emb2"

/lab_host="DH10B"

/note="vector: pCMV-Sport6"

1..2626

/gene="MGC52827"

/db_xref="GeneID:379106"

96..1592

/gene="MGC52827"

/codon_start=1

/product="MGC52827 protein"

/protein_id="AAH41742.1"

/db_xref="GI:27469877"

/db_xref="GeneID:379106"

/translation="MRVLEPPTEDTLIGFOAFMVYFSGGLSLAFIRGSEPS
FDYETTHVYSLSLVNSVLPSTEDTLIPCRPSVLYNPISVTSQDPLTEIH
NKNPVFGGERKPNCEKRNNTAIIIPRRETHRLHYLHFLPQOAHYIYI
IHQAGNSTFNRAKLNVGKEALRDEMDCLFLHDVLIPEMDFLVYCDPSPGHAS
VAMKFSYSLPYMPTFGVSALTPOYKMGFNPBYMGWGEDDDIATRYALGSMKI
TPSPVGVHYKMKVAKGQGNENPHREPLIRTOBMKVIDMNSLNTKLARELEPL
YTNVVDIGVDPRAKALRNRPPTPAISHYINTKMPADGOINPAPDHPNDSDIN
EAEISAEPRNKPVVAPSTTAEKLSIMPASA PHSAGEDKOISVIRBMKNNGRVK
EVOARVKKKSIVQSSKLEROVELFGARERPQDVGRSLPHLSAPL"

324..1130

/note="MGC52827"

/gene="Galactosyl T 2; Region: Galactosyltransferase. This
is a family of galactosyltransferases from a wide range of
Metazoa with three related galactosyltransferases
activities"

/db_xref="CD:DFam02709"

misc_feature

324..1130

/note="MGC52827"

/gene="Galactosyl T 2; Region: Galactosyltransferase. This
is a family of galactosyltransferases from a wide range of
Metazoa with three related galactosyltransferases
activities"

/db_xref="CD:DFam02709"

ORIGIN

Alignment Scores:

Pred. No.: 9,58e-58 Length: 2626
Score: 730.00 Matches: 145
Percent Similarity: 60.61% Conservative: 55
Best Local Similarity: 43.94% Mismatches: 120
Query Match: 35.94% Indels: 10
DB: 5 Gaps: 5

US-10-661-430-1 (1-383) x BC041742 (1-2626)

QY 55 VALLEUGLYASNTThAlaserThrSerAaPLeuAspThrTrpAsnSerThrPhe 74
DB 201 ATCTTCGAGCGCTCTCTGAGCCAGTTTGACTATACCGGACC---CATATGCTTAC 257
QY 75 SerProIleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleuPhe 94
DB 258 ACCAACCTCTCCCTGCTCAAC-----TTCATGCTGCCCTCTACT 299
QY 95 ProAspAsnGlnThrIleuGlnPheCysAsnGlnThrProProHISleuValGlyProIle 114
DB 300 GAAGAGGACACCTTACTTCTGATGCCCCACAGAGATCCCTTATTTGGTGGCCCCCAT 359
QY 115 ArgVal---PheIleuAspGluProAspPheIleuGluValIleTyxProAspThr 133

DB 360 TCCGTCACCTTCAGCAATCCCAATCCCACTGAGAGAGATTTCACACAGACCCCTCGTG 419
QY 134 HISLHAGLYGHIHISGLYMetProIleAspCysValAlaArgHISArgValAlaIleIle 153
DB 420 TTCAAAGGTGGAGATACCAAGCCCCCAATCTGTAGAGCCCGGACACACAGCCATCATC 479
QY 154 ValProIleArgAspArgGluAlaHISleuValrgIleMetLeuHISAsnLeuHISerLeu 173
DB 480 ATCCCTCACCGGACAGAGACATCTGAGGACCTTGCTCTTATCTGACCCCTTC 539
QY 174 LeuAlaIleGlnGlnIleuAspIleAlaIlePheIleValGlnGlnValAlaAsnGlnThr 193
DB 540 CTGCAGCCGACAGACCTCCTACAGATCTCATCATCATCATCATCATCATCATCATCAT 599
QY 194 PheAsnArgGlyIleuValMetAsnValGlyIleAspValAlaSerArgLeuTyxProThr 213
DB 600 TTCACCCGGGCGCAAGCTGTAAATGTTGGGTGAAGAGAGCTTGGGGATGAGGACTGG 659
QY 214 GlnCysPheIlePheHISAspValAspLeuIleProGluAspAspArgAsnLeuTyxThr 233
DB 660 GACTCTCTGTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 719
QY 234 Cys---ProIleGlnProArgHISMetSerValAlaIleAspIlePheAsnTyxIleu 252
DB 720 TGTGACCCCTGAGGCCCAACAGCCCTGTAGCCATGAACAAATTCAGCTACAGCCTG 779
QY 253 ProTyxSerAlaIlePheGlyIleSerAlaLeuTyxIleAspHISleuValIle 272
DB 760 CCATACCTTATGTTTGGGGAGTCTCTGCTCCCTTACCCCAACAGATATGTAAGATG 839
QY 273 AsnGlyPheSerAsnAspPheTrpGlyTyxPglyGlyGluAspAspLeuAlaThrArg 292
DB 840 AATGATTTCTTATGATATTTGGGAGAGGAGAGATGATGATGATGATGATGATGATG 899
QY 293 ThrSerMetAlaGlyLeuIleValSerArgTyxProThrGlnIleAlaArgTyxMet 312
DB 900 GTCCCTCTGGCGGATGATAAATATCACCCGCTCCGCTGCTGAGGCGCATATTAAGATG 959
QY 313 IleIleValSerThrGluAlaThrAsnProValAsnIleCysArgTyxIleIleMetGly 332
DB 960 GTACAGCAACAGGAGACAGGACAGGACAGGAGAAATCCACACAGCTTGTATATC 1019
QY 333 GlnThrIleArgArgTyxTrpArgAspGlyLeuSerAsnLeuTyxIleValAsn 352
DB 1020 AGGACCCAAAGATGTGAAAGTGACGCGCATGAATCTGCTGAACTACAGCTCTGGCC 1079
QY 353 LeuGlnLeuTyxProLeuTyxThrArgAlaValIleAspLeu---LeuGlnIleAspCys 371
DB 1080 AGAGAGCTAAGAACTCTCTTACACCAATGTGACCGTGATATAGGGGTGACCTTAGGCA 1139
QY 372 ArgArgGluLeuArgArgAspPheProThr 381
DB 1140 AAGAAAGCCCTCAGAAATCGGGGTCAAC 1169
RESULT 10
LOCUS BC075452 2340 bp mRNA linear VRT 24-NOV-2004
DEFINITION Xenopus tropicalis UDP-Gal:betaGal:GlcNAc beta 1,4-
galactosyltransferase, complete cds.
IMAGE:70117963), polypeptide 3, mRNA (CDNA clone MGC:89236
BC075452
BC075452.1 GI:49522603
KEYWORDS MGC.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huijck, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 2340)
Klein, S. and Gerhard, D.S.
Direct Submission
Submitted (28-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: xgc@dcfemail.nih.gov
Tissue Procurement: Rob Granger, University of Virginia
cDNA Library Preparation: Open Biosystems
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butlerfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu, Parvaneh Saeedi, J.R. Santos, Angélique Scherch, Ursula Skalska, Duane Smalins, Jeff Scott, Miranda Tsai, George Yang, Jacqui Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: <http://image.lnl.gov>
Series: IRAX Plate: 171 Row: P Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
Location/Qualifiers
1. 2340
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:89236 IMAGE:7017963"
/tissue_type="Whole body, male, 10 months old, strain F6, normal"
/clone_lib="NICHD_XGC_Swb1"
/lab_host="DH10B"
/note="Vector: pExpress1"
1. 2340
/gene="b4gal3-prov"
/note="synonym: MGC89236"
/db_xref="GeneID:448369"
227..1291
/gene="b4gal3-prov"
/codon_start=1
/product="UDP-Gal:betaglcNAc beta 1,4-galactosyltransferase, polypeptide 3"
/protein_id="AAH75452.1"
/db_xref="GI:49522604"
/db_xref="GeneID:448369"

/translation="MFARIDRRKFLLLCFSQLLELLYRRGAGIFQGLFGSDP
EPDYRTHDVTNLSLPVDDPENVASQCPASPIVGLSVSFNRPPLKVOQNR
YKPGYSRRCRFRYRAV.IPYRNEPHRLTYLHAPLQGLYAVFYQA
NGGTNRKALIKLDDMCILHDVDLIPENDYNIQDEEYKILASMD
KEDYSLPWTYINGISALTPDHYMLINGIPNNYKMGEDDIIARIIRAGSIARTP
LSVGYKMSIHRDSEYNSORVYDQLSVTRQTRKEDGMSLDPLKLSRTKALTYNI
TYDIDGEPPLVTEKPGMWLF"

ORIGIN

Alignment Scores:

Pred. No.:	1,036-57	Length:	2340
Score:	729.00	Matches:	145
Percent Similarity:	57.26%	Conservative:	64
Best Local Similarity:	39.73%	Mismatches:	122
Query Match:	35.89%	Indels:	34
DB:	5	Gaps:	5

US-10-661-430-1 (1-383) x BC075452 (1-2340)

Qy	8	ValAlaArgLeuLysSerLeuValLeuCybAlaValLeuLeuValAlaMet	27
		:::	:::
Db	242	ATGACCGGCGGAAGTTTCTGCTGCTTTGCTTCCAGCTGCTTCCATTAAATC	301
Qy	28	lLeTyrLys-----lLeProSerLeuTyrGluAsnLeuThrllLeGlySerSerThrLeu	45
		:::	:::
Db	302	CTGTACCGGAGAGGAGCTCCGGAATATTCAGGGGCTGTTGGGGGCTCCGAC-----	355
Qy	46	lLeAlaAspValAspAlaMetGluAlaValLeuGluYAsnThrAlaSerThrSerAspAsp	65
		-----	-----
Db	355	-----	-----
Qy	66	LeuLeuAspThrThrAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPhe	85
		:::	:::
Db	356	-----CCGGCTTGGAGTACTCCAGAACCCAGCGTATCACTTAACCTGATGTT	409
Qy	86	MetGluAspThrArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCybAsnGln	105
		:::	:::
Db	410	GTGCTGAC-----CCGGAATGTGAATTCAGATATGCGCAGCA	451
Qy	106	ThrProProHleuValGlyProIleArgVal---PheLeuAspGluProAspPheLys	124
		:::	:::
Db	452	AAATGCGCCATTGTTGGTGTCTCTGAGCGTTTCATCCACAACCCCGACCTGMAA	511
Qy	125	ThrLeuGluLyslLeTyrProAspThrHlSalGlylYHleGlyMetProLysAspCyb	144
		:::	:::
Db	512	AAGTGCACAGAAAAACCGCTACGTAAGCCGGGGGCTATTGACCCGAGACACTGC	571
Qy	145	ValAlaArgHleArgValAlaIleLeuValProTyrArgAspArgGluAlaHleuArg	164
		:::	:::
Db	572	TTCGGCGCTACGAAACAGACATTAATACCTACCGGAACGAAAGCGCACCTCCGC	631
Qy	165	lLeMetLeuHleAsnLeuHleSerLeuLeuAlaValGlnGlnLeuAspTyrAlaIlePhe	184
		:::	:::
Db	632	ACCTCTCTTACTACTGACACCCCTTCCTACGCGCAGACGCTGACCTGACCTTC	651
Qy	185	lLeValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyr	204
		:::	:::
Db	692	ATTGTCATCAGGCGCGGAACCGCACATTCATCGAACCAATCCCTAAACATTGGGGTG	751
Qy	205	AspValAlaSerArgLeuTyrProTTPGlnCybPheHlePheHleAspValAspLeu	224
		:::	:::
Db	752	AGGAGGCGCTAAACCTGACGACGTGACCTGCTGCTGACGACGCTGACCTGCTG	811
Qy	225	ProGluAspAspArgAsnLeuTyrThrCybProIleGln---ProArgHleMetSerVal	243
		:::	:::
Db	812	CCCGAGAACATTAACAACCTGTCATTTGGAGTAGAGAGATATCCCAAGCACTTAGCCAGC	871
Qy	244	AlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlylYlSerAla	263
		:::	:::
Db	872	GCATGACAAAGTTCCATGACCTGCGGTATGGACGTAATTTGAGAGGTGTGTGCA	931
Qy	264	LeuThrLysAspHleLeuLysLyslLeuGlnGlyPheSerAsnAspPheTTPGlyTTPGly	283
		:::	:::

Db 932 CTCACCCAGATCATACAGATGACCAAGCATCCCAACATTAATCTGGGGCTGGGGC 991
Qy 284 G|Y|G|U|A|S|P|A|S|P|L|E|U|A|L|A|T|H|A|T|A|T|S|E|R|M|E|T|A|G|L|Y|L|E|U|Y|V|A|L|S|E|R|A|T|G|T|Y|R 303
Db 992 GGAAGAGAGATGACATCTGATGAGATCCCGCTTACCGGAGATGAGCATTCCTCGGACC 1051
Qy 304 P|T|H|I|N|G|I|N|L|A|A|A|T|G|T|Y|R|Y|S|E|T|L|E|U|Y|S|H|I|S|E|R|T|H|L|A|L|A|T|H|A|S|P|R|O|V|A|L 323
Db 1052 CCAC|T|C|A|G|C|G|T|C|G|G|C|G|G|C|G|A|T|A|A|A|T|G|A|T|C|T|C|C|A|G|C|G|G|G|A|C|T|C|G|G|G|A|C|G|G|T|G 1111
Qy 324 A|S|N|L|Y|S|Y|A|T|G|T|Y|R|Y|S|I|L|E|T|E|G|Y|G|I|N|T|H|L|Y|S|A|T|G|T|T|H|A|T|G|A|S|P|G|I|Y|L|E|U 343
Db 1112 A|A|C|A|G|T|C|A|A|A|A|T|A|C|C|A|C|A|C|A|C|A|C|C|C|G|G|C|A|A|C|A|A|C|T|G|A|G|G|G|A|G|A|C|G|G|G|A|T|G 1171
Qy 344 S|E|R|A|N|L|E|U|Y|T|Y|R|Y|S|E|U|V|A|L|S|E|U|G|I|N|L|E|U|Y|P|R|O|L|E|U|Y|T|H|A|T|G|A|L|A|V|A|L 363
Db 1172 A|A|C|T|G|C|T|G|A|C|T|T|C|A|A|C|T|G|C|T|C|C|A|G|A|C|C|A|A|G|C|C|G|C|A|T|T|A|T|A|C|A|T|A|T|C|A|T 1231
Qy 364 V|A|L|S|P|L|E|U|L|E|U|G|I|N 368
Db 1232 G|T|G|A|T|A|T|C|G|G|G|A|G 1246
RESULT 11
AF142671 1952 bp mRNA linear ROD 01-JAN-2000
LOCUS Mus musculus beta-1,4-galactosyltransferase III mRNA, complete cds.
DEFINITION AF142671
ACCESSION AF142671 GI:6651181
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1952)
Lo,N.-W., Shaper,N.L. and Shaper,J.H.
MURINE BETA-1,4-GALACTOSYLTRANSFERASE FAMILY MEMBERS
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1952)
Lo,N.-W., Shaper,N.L. and Shaper,J.H.
DIRECT SUBMISSION
SUBMITTED (14-APR-1999) Oncology, Johns Hopkins, 600 N. Wolfe St.,
BALTIMORE, MD 21287, USA
FEATURES
source location/Qualifiers
1..1952
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
235..1422
/note="beta4GalT-III"
/codon_start=1
/product="beta-1,4-galactosyltransferase III"
/protein_id="AAF22221.1"
/db_xref="GI:6651182"
/translation="MLRLLEPCTLALVGSQALVMMWYLSIGFRSLSPALFGDPGP
TFDYSPHVDVYNSLHPALPAAGAPAPPAOLPYCPBSPFLVGVSPVSLAE
IYERHPVSGSRVAPAGEPSRTAIIYVPHARHNI,RLIYHAPLORQOLAVGI
YVTHQNGTFRACLIANGVREBARDEMDCLFLHVDLLPBNHNLTYCDPRGRH
VAVAANKFESLPYQYFGVSALTPTDQYFLKNGSPNEKVGSDDDIDATKVRLAGM
KLSRPTSVGHKWKHKGDKENBPHFDLVLTIONSWTQDNKNSLYRLAELG
PLSTPTITADIGTDPGRAPASGRYPVPGSSQAFROEMLQRRPPARPGLPTANHTAPR
GSH"

Qy 95 P|R|O|A|E|A|N|G|I|N|T|H|L|E|U|G|I|N|P|H|E|C|Y|S|A|N|G|I|N|T|H|R|P|R|O|H|I|E|L|E|U|V|A|L|Y|P|R|O|I|L|E 114
Db 445 C|C|T|C|G|G|G|C|T|C|A|A|G|C|T|T|G|C|C|T|A|C|T|G|T|C|C|A|A|A|A|T|C|A|C|C|T|T|T|A|T|G|T|G|G|T|C|T|G|G 504
Qy 115 A|T|G|V|A|L|---|P|H|E|U|S|A|S|P|L|U|P|R|O|A|S|P|H|E|Y|S|T|H|L|E|U|G|I|N|L|E|Y|L|E|T|Y|R|P|R|O|A|S|T|H| 133
Db 505 T|C|A|G|T|A|C|C|T|T|T|A|G|C|C|C|A|G|T|G|C|C|C|T|C|A|T|G|C|G|G|A|T|G|T|G|G|A|T|G|C|C|G|G|G|T|G 564
Qy 134 H|I|S|A|I|A|G|Y|G|I|N|H|I|G|I|Y|M|E|R|P|R|O|L|Y|S|A|S|P|C|Y|S|V|A|L|A|A|A|G|H|I|A|T|G|V|A|L|A|I|L|E|I|E 153
Db 555 G|A|A|T|C|A|G|G|G|G|C|G|G|T|A|C|G|T|C|T|G|C|A|G|G|G|T|G|A|G|C|T|C|G|C|C|G|C|C|G|A|C|G|C|A|A|T|A|T 624
Qy 154 V|A|L|P|R|O|T|Y|R|A|T|G|A|S|P|A|T|G|I|N|L|A|I|S|E|U|Y|R|G|I|L|E|T|E|L|E|U|N|H|I|A|S|P|L|E|U|H|I|S|E|R|L|E|U 173
Db 625 G|T|G|C|C|C|A|T|C|G|G|C|C|G|G|G|A|G|C|A|C|C|T|T|G|G|C|G|C|G|C|T|T|A|C|A|C|T|G|C|A|C|C|C|T|T|G 684
Qy 174 L|E|U|A|L|A|Y|S|G|I|N|L|E|U|S|A|S|P|T|Y|R|A|L|I|E|P|H|E|I|E|V|A|L|G|I|N|V|A|L|A|S|N|G|I|N|T|H|R 193
Db 685 C|T|G|C|A|G|C|C|C|A|G|C|A|C|T|T|G|C|T|A|C|G|C|A|T|T|A|T|G|C|A|T|C|C|A|C|A|G|C|T|G|A|A|A|T|G|A|A|C|G 744
Qy 194 P|H|E|A|N|A|T|G|Y|L|Y|S|L|E|U|M|E|R|A|S|N|V|A|L|G|I|Y|R|A|S|P|V|A|L|A|S|E|R|A|T|G|L|E|U|Y|R|P|R|O|T| 213
Db 745 T|T|T|A|C|A|G|G|C|A|A|A|C|T|G|T|G|A|A|C|T|A|G|G|G|T|G|A|G|G|A|A|G|C|C|T|T|C|T|G|A|T|G|A|A|A|T|G|G 804
Qy 214 G|I|N|C|Y|S|P|H|E|I|E|P|H|E|H|I|A|S|P|V|A|L|S|P|L|E|U|N|P|R|O|G|I|U|A|S|P|A|S|P|A|S|N|L|E|U|Y|R|T|H|R 233
Db 805 G|A|C|T|C|T|G|T|T|T|T|A|C|A|G|C|A|G|T|G|A|C|T|C|T|T|C|C|A|A|A|A|G|C|A|A|C|A|T|C|T|G|H|A|T|G|G 864
Qy 234 C|Y|S|---|P|R|O|I|L|E|I|N|P|R|O|A|S|H|I|S|E|R|V|A|L|A|I|E|A|S|P|L|Y|S|P|H|E|A|S|N|T|Y|R|Y|S|L|E|U 252
Db 865 T|G|C|A|G|C|C|C|C|G|G|G|A|C|C|C|G|C|A|C|G|T|G|C|T|G|T|G|C|A|T|G|A|A|A|A|G|T|T|G|A|T|A|C|A|G|C|C|T|C 924
Qy 253 P|T|O|T|Y|R|S|E|R|A|L|I|E|P|H|G|I|G|Y|L|E|S|E|R|A|L|E|U|N|T|H|L|Y|S|A|P|H|I|S|L|E|U|Y|S|I|L|E 272
Db 925 C|C|G|T|A|C|C|C|C|C|A|G|T|A|C|T|T|G|G|G|A|G|T|T|C|A|G|C|G|C|A|C|C|C|G|C|C|G|A|C|A|G|T|A|C|T|G|A|A|G|A|T|G 984
Qy 273 A|S|N|G|I|P|H|E|S|E|R|A|S|P|A|S|P|P|H|E|T|R|O|L|Y|R|P|G|I|Y|G|I|Y|G|I|U|A|S|P|A|S|P|L|E|U|A|L|A|T|H|A|T|A|T|G 292
Db 985 A|A|G|G|C|T|T|C|C|C|A|G|A|G|A|T|A|C|T|G|G|G|G|C|G|G|G|G|G|C|G|A|G|A|T|G|A|C|A|T|T|C|T|A|C|C|A|G 1044
Qy 293 T|H|S|E|R|M|E|T|A|I|G|L|Y|L|E|U|Y|S|E|R|A|T|G|T|Y|R|P|R|O|H|I|N|G|I|N|L|A|I|A|T|G|T|Y|R|Y|S|E|T 312
Db 1045 G|T|C|G|C|C|T|G|G|C|T|G|G|G|A|T|G|A|G|A|T|T|C|T|G|C|A|C|A|C|A|C|T|C|T|G|G|G|A|C|A|T|G|A|A|G|A|T|G 1104
Qy 313 I|L|E|Y|S|H|I|S|E|R|T|H|L|A|L|A|T|H|A|S|P|R|O|V|A|L|S|E|U|Y|S|Y|A|T|G|T|Y|S|I|L|E|T|E|G|Y 332
Db 1105 G|T|G|A|G|C|A|C|A|G|G|G|A|T|A|A|G|A|A|A|G|A|A|A|C|C|C|C|A|C|A|G|A|T|T|G|A|C|C|T|G|G|T|C 1164
Qy 333 G|I|N|T|H|L|Y|S|A|T|G|T|T|H|A|T|A|T|S|E|R|A|N|L|E|U|Y|T|Y|R|Y|S|E|U|V|A|L|S|N 352
Db 1165 C|G|T|A|C|C|C|A|A|A|T|T|C|T|T|G|A|C|A|C|A|G|A|T|G|A|A|T|G|A|C|T|A|C|T|A|C|G|A|C|G|A|C|T|G|G|C|A 1224
Qy 353 L|E|U|G|I|N|L|E|U|Y|P|R|O|L|E|U|Y|T|H|A|T|G|A|L|A|V|A|L|S|P|L|E|U 366
Db 1225 A|G|A|G|A|G|C|T|G|G|G|C|C|T|C|T|A|T|A|C|A|A|C|A|T|G|C|A|G|A|C|A|T|C 1266
RESULT 12
BC013619 1962 bp mRNA linear ROD 04-OCT-2003
LOCUS Mus musculus UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase,
polypeptide 3, mRNA (cDNA clone MGC:11711 IMAGE:396561), complete
cds
ACCESSION BC013619
VERSION BC013619.1 GI:15488995
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1962)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

US-10-661-430-1 (1-383) x AF142671 (1-1952)

DB 1138 GTGAAGCAGAGGAGGATTAAGAAATGAGAAATCCCAAGATTGACCTCTGGCC 1197
QY 333 GlnThrLysArgArgTrrThrArgAspGlyLeuSerAsnLeuTyrrLysLeuValasn 352
DB 1198 GGTACCCAGAAATTTCTTGACACAGATGAAATGAATCACTACGACCGACTGCGCA 1257
QY 353 LeuGluLeuLysProLeuTyrrThrArgAlaValAlaAspLeu 366
DB 1258 AGAGAGCTGGGCTCTCTCTATACCAACATCACTGACGACATC 1299
RESULT 13
LOCUS BT020682
DEFINITION Bos taurus UDP-Gal:betaGalNAc beta 1,4- galactosyltransferase,
BT020682
ACCESSION BT020682 1927 bp mRNA linear MAM 16-FEB-2005
VERSION BT020682.1 GI:59857728
KEYWORDS FLI_CDNA.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1927)
AUTHORS Smith,T.P., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Bennett,G.L., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Chitko-McKown,C.G., Perce,G., Holt,I., Karaymicheva,S., Liang,F.,
Quackenbush,J. and Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
JOURNAL 11282978
PUBMED 2 (bases 1 to 1927)
REFERENCE Hathay,G.P., Sonstegard,T.S., Clawson,M.L., Heaton,M.P.,
AUTHORS Keefe,J.W., Snelling,W.M., Weidmann,R.T. and Smith,T.P.L.
Sequencing and analysis of Bos taurus full-length insert cDNA
clones
TITLE Unpublished
JOURNAL 3 (bases 1 to 1927)
REFERENCE Hathay,G.P., Sonstegard,T.S., Clawson,M.L., Heaton,M.P.,
AUTHORS Keefe,J.W., Snelling,W.M., Weidmann,R.T. and Smith,T.P.L.
Direct Submission
TITLE Submitted (16-FEB-2005) Molecular Genetic Research Unit, USDA-ARS
JOURNAL U.S. Meat Animal Research Center, P.O. Box 166, Spur 18D, Clay
Contact: Gregory P. Hathay at hathay@email.marc.usda.gov.
COMMENT location/Qualifiers
FEATURES
source 1. 1927
/organism="Bos taurus"
/mol_type="RNA"
/db_xref="taxon:9913"
/clone="1BOV28F7; plate 28 row F column 7"
/cissue_type="pooled"
/clone_id="1BOV"
/lab_host="Escherichia coli DH10B"
/note="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
1. 1927
/gene="B4GALT3"
/note="putative, similar to GeneID: 8703"
221. 1411
/gene="B4GALT3"
/codon_start=1
/product="UDP-Gal:betaGalNAc beta 1,4-
galactosyltransferase 3"
/protein_id="AA008699.1"
/db_xref="GI:59857728"
/translation="MLRILRPCTLLIVSGLAVMYTSLGGRSLAFGRBOEP
AFVSPHIDVYNSLHMPGAPVAPGGLPAPGGLPYCPSPSLVLPISVSPVSLA
EIVENRIVEPGRYRPARCEPRRTAIIVPRARENHRLYLHLAPFLPOQLAYG
IYVHQAGNGTFNRAKLINVGVEALRDEMDCLFLHVDLLPNDHNLVYCDPRGR

ORIGIN
HVAAMNKFYSLPDYFGVSALTPDQILNMNGFPNIEYWGGEEDDIATRVLAG
MKISRPTSVGHKRVKRGDKNBENRFDLIVRTONSMTQGMNLVTYOLSLREL
GLXTNITADIGTPDPRGRPTSSGHHYPPGSSGARQENLQRRPARRPPRLTAHTAP
HGSH"
Alignment Scores:
Pred. No.: 4.51e-57 Length: 1927
Score: 721.00 Matches: 136
Percent Similarity: 64.29% Conservative: 44
Best Local Similarity: 48.57% Mismatches: 98
Query Match: 35.50% Indels: 2
DB: 4 Gaps: 2
US-10-661-430-1 (1-383) x BT020682 (1-1927)
QY 89 lIeaRgProLleuPheProAspAsnGlnThrLeuGlnPheCyAsnGlnThrProPro 108
DB 416 GTTGCCCGAGGGGGCTCCAGACTCCCAAGGCTGCATCTGCGCAAGATCTCCT 475
QY 109 HisLeuValAlaIleIleValProTyrrArgAspArgIuaIaHisLeuArgIleMetLeu 127
DB 476 CTCTTAGTGGGTCCCAATCCGTCCTTTAGCCCAAGTCCATCGTCGACAGATTGTG 535
QY 128 LysIleTyrrProAspThrHisAlaGlyIleHisGlyMetProLysAspCyLeuAlaIaArg 147
DB 536 GAGAGGAATCCCGGGGTGAACCCGGGGGGCCGATACCCGCTTCGAAGTGTGAGCCCGG 595
QY 148 HisArgValAlaIleIleValProTyrrArgAspArgIuaIaHisLeuArgIleMetLeu 167
DB 596 TCCCGGACAGCATCATTTGTGCGCCACCGTCCGGGAGACACACTCGCTGCTGCTC 655
QY 168 HisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrrAlaIlePheIleValGlu 187
DB 656 TACCACTGCACCCGCTTCCTCCAGCGCAGACGCTGTTCAGGATTAATGTCATCAC 715
QY 188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrrAspValAla 207
DB 716 CAGCTGGAAATGAAACATTTTAACAGGGCAACGCTCGAATGTGCGGTGGCGGAGGCC 775
QY 208 SerArgLeuTyrrProTrrGlnCysPheIlePheHisAspValAspLeuProGluAsp 227
DB 776 CTGCGAGATGAAGATGGAGCTGCTTGTTCGATGATGTGATCTCTCGCAGAGAAC 835
QY 228 AsparGlnLeuTyrrThrCys---ProIleGlnProAlaGlyMetSerValAlaIleAsp 246
DB 836 GACCACAACTGTATGTGTGGCATCCCGAGACCTCGGACGCTTGCTGTGCCATGAAC 895
QY 247 LysPheAsnTyrrLysLeuProTyrrSerAlaIlePheGlyIleSerAlaLeuThrLys 266
DB 896 AATTTGGATACAGCTCCCGTATCCCAAGTACTTTGGCGGGGTCTCAGCGCTCCTCT 955
QY 267 AspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTrrGlyTyrrGlyIleAsp 286
DB 956 GACCAGTACCTGAAGATGATGGCTTCCCAATGAATACTGGGGGTGGGGTGTGAGGAT 1015
QY 287 AsparLeuAlaIthrArgThrSerMetAlaGlyLeuLysValSerArgTyrrProThrGln 306
DB 1016 GACGCACTGTGTACCAAGGTTGCTGGCTGGGAGGAATATCTGCTCCGCCCATCTCC 1075
QY 307 lIeaIaArgTyrrLysMetIleLysHisSerThrgluAlaIthrAsnProValAsnLysCys 326
DB 1076 GTGGACACTTATAGATGTGAAGCACCGAGAGATTAAGGCGACGAGAAAATCCCAAC 1135
QY 327 ArgTyrrLysIleMetGlyGlnThrLysArgArgTrrThrArgAspGlyLeuSerAsnLeu 346
DB 1136 AGGTTAGACCTTCGTGCGTACCAAGCAATCTCGACTCAAGATGAGATGAATCACTG 1195
QY 347 LysTyrrLysLeuValAsnLeuGlnLeuLysProLeuTyrrThrArgAlaValAlaAspLeu 366
DB 1196 ACATACAGTTGTGCTGTGAGAGCTGGCCCTCTCTATACGAACATCAGACAGACATT 1255
RESULT 14

BC087436 1329 bp mRNA linear VRT 01-FEB-2005
LOCUS BC087436
DEFINITION Xenopus laevis hypothetical LOC496041, mRNA (cDNA clone MGC:99248
IMAGE:7210136), complete cds.
ACCESSION BC087436
VERSION BC087436.1 GI:56270475
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 1329)
AUTHORS Klein S.L., Strauberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 1329)
AUTHORS Strauberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Teshiyuki, S.,
Carninci, P., Prange, C., Kana, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 3 (bases 1 to 1329)
AUTHORS Klein, S. and Gerhard, D.S.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
REMARK NIH-MGC Project
COMMENT Contact: XGC help desk
Email: xgcs-help@nih.gov
Tissue Procurement: Igor B. David
cDNA Library Preparation: Express Genomics
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
FEATURES
Source
1..1329
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"

/clone="MGC:99248 IMAGE:7210136"
/cise_type="Testis, adult male, wild type"
/clone_id="NICHD_XGC_Te2"
/lab_host="DH10B"
/note="Vector: pExpress1"
1..1329
/gene="LOC496041"
/db_xref="GeneID:496041"
190..1254
/gene="LOC496041"
/codon_start=1
/product="LOC496041 protein"
/protein_id="AAH87436.1"
/db_xref="GI:56270476"
/db_xref="GeneID:496041"
/translation="MPLSRIDKGNLILLCFSQLFLILYRGAGVFGQGFSGSD
PMDYSKIDVYTNLSVFPNPENVTQYCPASPLIVGPLSVSPQPTPLKYOQNR
YKPGGYFSPRHCFGRYRTAVIIPHNRRHARLTLLYLHPFLQQLHALFVHQ
GNSFRRAKLINIGVREALKDEWDCLIHVDVLPENDYNDYICDEEYKILASMD
KEHYSLPYWTYFGVSALPDPQYMRNGFRPNISYMGSGEDDILAMRILAGMSITRP
LSIGRYKMTSHNRDSNGEENSRYDQLGRTTRTMEDEKNSLDKILSRTRAPLYTNI
TYAIGDAPLVETTPAQMKLF"

ORIGIN
Alignment Scores:
Pred. No.: 3,54e-57 Length: 1329
Score: 720.00 Matches: 143
Percent Similarity: 57.46% Conservative: 61
Best Local Similarity: 40.28% Mismatches: 117
Query Match: 35.45% Indels: 34
Gaps: 5
US-10-661-430-1 (1-383) x BC087436 (1-1329)
QY 14 LeuLeuValLeuCyAlaValLeuLeuValHisAlaMetIleTyr-----Ile 31
Db 223 CTGCTCTGCTGCTGCTTCTCTGCTGCTTCTTCTTAAATCTATACGAGAGGGCT 282
QY 32 ProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIleAlaValaPa 51
Db 283 GCCGCGCTCTTCAGAGCGCTGTTGGGGATCCGAT----- 318
QY 52 MetGluAlaValLeuGlyAsnThrAlaSerThrSerAspAspLeuAspThrTPa 71
Db 319 -----CCGCGCTTGGAC 330
QY 72 SerThrPheSerProIleSerGluValAsnGlnThrSerPheMetGluAspIleA 91
Db 331 TACTCCAAACCATGATGATGATACCAACCTGAGTGTGTGTCACCA----- 381
QY 92 IleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeu 111
Db 382 -----CCAGAGAAATGTAATACTCACTATGCCCAGCAAAATCACCATACTG 432
QY 112 GlyProIleArgVal---PheLeuAspGluProAspPheLeuThrLeuGluIly 130
Db 433 GGTCTCTTATAGTTCATTCATTCGAAACCCCGACCTGAAAGAGTGCAGCAAAAAC 492
QY 131 ProAspThrHisAlaGlyIleGlyMetProIlyAspCyAlaAlaArgHisArgVal 150
Db 493 CGATATGTGAAGCGCTGCGGCTTCTTCAAGCCGACGACCTGTTGGCGCTACAGAAC 552
QY 151 AlaIleIleValProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnLeu 170
Db 553 GCCGCTATTAATCTACACGAAACAGAGAGCGCACCTCGCACCTCTTACTACTG 612
QY 171 HisSerLeuLeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGlnVala 190
Db 613 CACCCCTTCTTCAAGGCCAGAGCTTCAATATGACCTTCAATGTCACAGCGCGG 672
QY 191 AsnGlnThrPheAsnArgGlyIlyLeuMetAsnValGlyTyrAspValAlaSerArgLeu 210
Db 673 AACAGACATTCATGAGCGCAAACTGTAACATTTGGGGGTGAGGAGGCCCTAAAC 732

Qy		211	TyrProTrpGlnCysPheIlePheHisAspValAspLeuLeuProGluAspArgAsn	230
Dd		733	GACAGTGGAGCTCCCGATCTTTACAGACGTCGACTGGTCCCTGGAAACGATTACAC	792
Qy		231	LeuTyThrCysProIleGln---ProArgHisMetSerValAlaIleAspLysPheAsn	249
Dd		793	TTGTGCATTTTGTTGATGAGAGTAATGCCAACACCTGGCCAGCGCATGAGCAATTCCAC	852
Qy		250	TyrLysLeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuTyrLysAspHisLeu	269
Dd		853	TACAGCTTGGCGGTGGACCGTACTTTGGAGAGAGTGCTGCACATCACCCCAGATCAGTAC	912
Qy		270	LysLysIleAsnGlyPheSerAsnAspPheTrpGlyTyrProGlyGlyLysAspAspLeu	289
Dd		913	ATGGAGATCAATGGAATTCGCCAACAGTTACTGGGGTGGGGCGGAGAGATGATGACATC	972
Qy		290	AlaThrArgTrpSerMetAlaGlyLeuLysValSerArgTyrProTrpGlnIleAlaArg	309
Dd		973	GCTATGAGAGATCCGCTCAGCGGAGATAGCATCACCCGAGACCCTTCAGCTCGGCCGC	1032
Qy		310	TyrLysMetIleLysHisSerThrGluAlaThrAnProValaLeuLysCysArgTyrLys	329
Dd		1033	TATAAATGATCTCCCATTAACCGGAGCTCCGGCAACGAGAGACAGAAAAGTTAGAC	1092
Qy		330	IleMetGlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLys	349
Dd		1093	CMACTCGGCAACACCCGGAGAACAATGAGAGACAGACGGAGAACTCGCTGCATTCMAA	1152
Qy		350	LeuValaLeuLeuGluLeuLysProLeuTyrThrArgAlaValaVal	364
Dd		1153	CTGATCTCCAGACACGAGGCCCCCTCTATTCCAAATCACTACTGG	1197
RESULT 15				
LOCUS		BT020916	1921 bp	mRNA linear MAM 16-FEB-2005
DEFINITION		Bos taurus UDP-gal-4-epimerase beta 1,4-galactoseyltransferase,		
VERSION		polypeptide 3 (B4GALT3), mRNA, complete cds.		
KEYWORDS		BT020916		
SOURCE		BT020916.1 GI:59858196		
ORGANISM		FIL CDNA.		
		Bos taurus (cow)		
REFERENCE		Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;		
		Pecora; Bovidae; Bovinae; Bos.		
		Smith,T.P., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,		
		Caesa,E., Wray,J.E., White,J., Cho,D., Fahrenkrug,S.C.,		
		Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,		
		Chitko-Mckown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,		
		Quackenbush,J. and Keefe,J.W.		
		Sequence evaluation of four pooled-tissue normalized bovine cDNA		
		libraries and construction of a gene index for cattle		
		Genome Res. 11 (4), 626-630 (2001)		
JOURNAL		BT020916		
PUBMED		11282978		
REFERENCE		2 (bases 1 to 1921)		
AUTHORS		Hathey,G.P., Sonstegard,T.S., Clawson,M.L., Heaton,M.P.,		
		Keefe,J.W., Snelling,W.M., Weidmann,R.T. and Smith,T.P.L.		
		Sequencing and analysis of Bos taurus full-length insert cDNA		
		clones		
TITLE		Unpublished		
JOURNAL		3 (bases 1 to 1921)		
REFERENCE		Hathey,G.P., Sonstegard,T.S., Clawson,M.L., Heaton,M.P.,		
AUTHORS		Keefe,J.W., Snelling,W.M., Weidmann,R.T. and Smith,T.P.L.		
		Direct Submission		
TITLE		Submitted (16-FEB-2005) Molecular Genetic Research Unit, USDA-ARS		
JOURNAL		U.S. Meat Animal Research Center, P.O. Box 166, Spur 18D, Clay		
		Center, NE 68913		
COMMENT		Contact: Gregory P. Hathey at hathey@email.marc.usda.gov.		
FEATURES		location/qualifiers		
SOURCE		1..1921		
		/organism="Bos taurus"		

gene	CDs	ORIGIN	Alignment Scores:	Best Local Similarity:	Best Match:
gene	CDs	ORIGIN	Alignment Scores:	Best Local Similarity:	Best Match:
1. .1921	1. .1921	1. .1921	1. .1921	1. .1921	1. .1921
/gene="B4GL13"	/gene="B4GL13"	/gene="B4GL13"	/gene="B4GL13"	/gene="B4GL13"	/gene="B4GL13"
/note="putative, similar to GenID: 8703"	/note="putative, similar to GenID: 8703"	/note="putative, similar to GenID: 8703"	/note="putative, similar to GenID: 8703"	/note="putative, similar to GenID: 8703"	/note="putative, similar to GenID: 8703"
215. .1405	215. .1405	215. .1405	215. .1405	215. .1405	215. .1405
/codon_start=1	/codon_start=1	/codon_start=1	/codon_start=1	/codon_start=1	/codon_start=1
/product="UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase 3"	/product="UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase 3"	/product="UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase 3"	/product="UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase 3"	/product="UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase 3"	/product="UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase 3"
/protein_id="AA08933.1"	/protein_id="AA08933.1"	/protein_id="AA08933.1"	/protein_id="AA08933.1"	/protein_id="AA08933.1"	/protein_id="AA08933.1"
/db_xref="GI:59858197"	/db_xref="GI:59858197"	/db_xref="GI:59858197"	/db_xref="GI:59858197"	/db_xref="GI:59858197"	/db_xref="GI:59858197"
/translation="MLRRLRERPTLLALLVGSQALVVMYLSLGSFRRISALFGRBOE AFDSDPHDYSLNHPGAPVAPAGGAPADGLPYCKRSLVGPISVSPVPSLS EYERNRVRPGGRYPARCEPRSRTAIYVRRARHEIRLLPHLPLOOOLAY IYVHOGNGTGFENAKLILNVRALDEBMDCLFHDVDFPNDNLVYCDRGRPG HVAANMKPGSLPYQVPGVSALTPDQYIKMGKFNENKMGKGGDDIATVRLAQ MKLSRPTSVGHYRMVAKRGDKGNEENPHRPDLVLRONSSTQGMNSLTQLSREI GPLSTNTTADIGTDPRGPRTSSGPHYPPGSSQARQEMLORPPARPGRPLETNNHTA HGSH"	/translation="MLRRLRERPTLLALLVGSQALVVMYLSLGSFRRISALFGRBOE AFDSDPHDYSLNHPGAPVAPAGGAPADGLPYCKRSLVGPISVSPVPSLS EYERNRVRPGGRYPARCEPRSRTAIYVRRARHEIRLLPHLPLOOOLAY IYVHOGNGTGFENAKLILNVRALDEBMDCLFHDVDFPNDNLVYCDRGRPG HVAANMKPGSLPYQVPGVSALTPDQYIKMGKFNENKMGKGGDDIATVRLAQ MKLSRPTSVGHYRMVAKRGDKGNEENPHRPDLVLRONSSTQGMNSLTQLSREI GPLSTNTTADIGTDPRGPRTSSGPHYPPGSSQARQEMLORPPARPGRPLETNNHTA HGSH"	/translation="MLRRLRERPTLLALLVGSQALVVMYLSLGSFRRISALFGRBOE AFDSDPHDYSLNHPGAPVAPAGGAPADGLPYCKRSLVGPISVSPVPSLS EYERNRVRPGGRYPARCEPRSRTAIYVRRARHEIRLLPHLPLOOOLAY IYVHOGNGTGFENAKLILNVRALDEBMDCLFHDVDFPNDNLVYCDRGRPG HVAANMKPGSLPYQVPGVSALTPDQYIKMGKFNENKMGKGGDDIATVRLAQ MKLSRPTSVGHYRMVAKRGDKGNEENPHRPDLVLRONSSTQGMNSLTQLSREI GPLSTNTTADIGTDPRGPRTSSGPHYPPGSSQARQEMLORPPARPGRPLETNNHTA HGSH"	/translation="MLRRLRERPTLLALLVGSQALVVMYLSLGSFRRISALFGRBOE AFDSDPHDYSLNHPGAPVAPAGGAPADGLPYCKRSLVGPISVSPVPSLS EYERNRVRPGGRYPARCEPRSRTAIYVRRARHEIRLLPHLPLOOOLAY IYVHOGNGTGFENAKLILNVRALDEBMDCLFHDVDFPNDNLVYCDRGRPG HVAANMKPGSLPYQVPGVSALTPDQYIKMGKFNENKMGKGGDDIATVRLAQ MKLSRPTSVGHYRMVAKRGDKGNEENPHRPDLVLRONSSTQGMNSLTQLSREI GPLSTNTTADIGTDPRGPRTSSGPHYPPGSSQARQEMLORPPARPGRPLETNNHTA HGSH"	/translation="MLRRLRERPTLLALLVGSQALVVMYLSLGSFRRISALFGRBOE AFDSDPHDYSLNHPGAPVAPAGGAPADGLPYCKRSLVGPISVSPVPSLS EYERNRVRPGGRYPARCEPRSRTAIYVRRARHEIRLLPHLPLOOOLAY IYVHOGNGTGFENAKLILNVRALDEBMDCLFHDVDFPNDNLVYCDRGRPG HVAANMKPGSLPYQVPGVSALTPDQYIKMGKFNENKMGKGGDDIATVRLAQ MKLSRPTSVGHYRMVAKRGDKGNEENPHRPDLVLRONSSTQGMNSLTQLSREI GPLSTNTTADIGTDPRGPRTSSGPHYPPGSSQARQEMLORPPARPGRPLETNNHTA HGSH"	/translation="MLRRLRERPTLLALLVGSQALVVMYLSLGSFRRISALFGRBOE AFDSDPHDYSLNHPGAPVAPAGGAPADGLPYCKRSLVGPISVSPVPSLS EYERNRVRPGGRYPARCEPRSRTAIYVRRARHEIRLLPHLPLOOOLAY IYVHOGNGTGFENAKLILNVRALDEBMDCLFHDVDFPNDNLVYCDRGRPG HVAANMKPGSLPYQVPGVSALTPDQYIKMGKFNENKMGKGGDDIATVRLAQ MKLSRPTSVGHYRMVAKRGDKGNEENPHRPDLVLRONSSTQGMNSLTQLSREI GPLSTNTTADIGTDPRGPRTSSGPHYPPGSSQARQEMLORPPARPGRPLETNNHTA HGSH"
US-10-661-430-1 (1-383) x BT020916 (1-1921)	US-10-661-430-1 (1-383) x BT020916 (1-1921)	US-10-661-430-1 (1-383) x BT020916 (1-1921)	US-10-661-430-1 (1-383) x BT020916 (1-1921)	US-10-661-430-1 (1-383) x BT020916 (1-1921)	US-10-661-430-1 (1-383) x BT020916 (1-1921)
89 IleaArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProPro 108	89 IleaArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProPro 108	89 IleaArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProPro 108	89 IleaArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProPro 108	89 IleaArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProPro 108	89 IleaArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProPro 108
410 GTTGGCCCAAGGGGGGCTCCAGCTCTCTAAGGCTGCATCTGTCCTCAAGATCTCT 469	410 GTTGGCCCAAGGGGGGCTCCAGCTCTCTAAGGCTGCATCTGTCCTCAAGATCTCT 469	410 GTTGGCCCAAGGGGGGCTCCAGCTCTCTAAGGCTGCATCTGTCCTCAAGATCTCT 469	410 GTTGGCCCAAGGGGGGCTCCAGCTCTCTAAGGCTGCATCTGTCCTCAAGATCTCT 469	410 GTTGGCCCAAGGGGGGCTCCAGCTCTCTAAGGCTGCATCTGTCCTCAAGATCTCT 469	410 GTTGGCCCAAGGGGGGCTCCAGCTCTCTAAGGCTGCATCTGTCCTCAAGATCTCT 469
109 HisLeuValGlyProIleArgVal--PheLeuAspGlnProAspPheLeuThrLeuGln 127	109 HisLeuValGlyProIleArgVal--PheLeuAspGlnProAspPheLeuThrLeuGln 127	109 HisLeuValGlyProIleArgVal--PheLeuAspGlnProAspPheLeuThrLeuGln 127	109 HisLeuValGlyProIleArgVal--PheLeuAspGlnProAspPheLeuThrLeuGln 127	109 HisLeuValGlyProIleArgVal--PheLeuAspGlnProAspPheLeuThrLeuGln 127	109 HisLeuValGlyProIleArgVal--PheLeuAspGlnProAspPheLeuThrLeuGln 127
470 CTCTTAGGGGTCCCATATCCGTGTCCTTAGCCCAAGCCATGCTGGCAGAGATTGTG 529	470 CTCTTAGGGGTCCCATATCCGTGTCCTTAGCCCAAGCCATGCTGGCAGAGATTGTG 529	470 CTCTTAGGGGTCCCATATCCGTGTCCTTAGCCCAAGCCATGCTGGCAGAGATTGTG 529	470 CTCTTAGGGGTCCCATATCCGTGTCCTTAGCCCAAGCCATGCTGGCAGAGATTGTG 529	470 CTCTTAGGGGTCCCATATCCGTGTCCTTAGCCCAAGCCATGCTGGCAGAGATTGTG 529	470 CTCTTAGGGGTCCCATATCCGTGTCCTTAGCCCAAGCCATGCTGGCAGAGATTGTG 529
128 LysIleLeuTrpProAspThrHisAlaGlyGlyHisIsoleuMetProLysAspCysValAlaArg 147	128 LysIleLeuTrpProAspThrHisAlaGlyGlyHisIsoleuMetProLysAspCysValAlaArg 147	128 LysIleLeuTrpProAspThrHisAlaGlyGlyHisIsoleuMetProLysAspCysValAlaArg 147	128 LysIleLeuTrpProAspThrHisAlaGlyGlyHisIsoleuMetProLysAspCysValAlaArg 147	128 LysIleLeuTrpProAspThrHisAlaGlyGlyHisIsoleuMetProLysAspCysValAlaArg 147	128 LysIleLeuTrpProAspThrHisAlaGlyGlyHisIsoleuMetProLysAspCysValAlaArg 147
530 GAGGAGATCCCGCGGTAAACACGAGGGGGCGGATACCGCTCGCAAGGTGTGAGACCCCGG 589					

```
Db      890  AAATTGGATACAGCCCTCCCGTATCCCACTACTTGGCGGGGCTCCAGGCTCACTCCT 949
Qy      267  AsPHISleuLysLysIleasnGlyPheSerAsnAspPheTrpGlyTyrGlyGluAsp 286
Db      950  GACCACTACCTGAAAGATGAATGGCTTCCCAATGAATATCGGGGCTGGGGTGAAGAT 1009
Qy      287  AspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGln 306
Db      1010 GACGACATTCTTACCCAGGCTTCCTGGCTGGGATGAAATATCTCCTCCCCCACAATCC 1069
Qy      307  IleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCys 326
Db      1070 GTGGGACACTATAGAAGTGAAGACACGAGAGATTAAGGCAACGAGAAATCCCCAC 1129
Qy      327  ArgTyrLysIleMetGlyGlnThrLysArgArgTrpThrArgAspGlyLeuSerAsnLeu 346
Db      1130 AGGTTTGACCTCCCTGGTCCGTACCCAGAAATCTCTGACTCAAGATGGAATGAATCACTG 1189
Qy      347  LysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAlaValValAspLeu 366
Db      1190 ACATACCAGTTGCTGCTCGAGAGCTGGGCCCTCTTATACGAACATCACAGCAGACATT 1249
```

Search completed: January 9, 2006, 03:00:01
Job time : 6403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 9, 2006, 00:36:56 : Search time 4931 Seconds
(without alignments)
3634.040 Million cell updates/sec

Title: US-10-661-430-1

Perfect score: 2031
Sequence: 1 MAFRLAVARLKSLLVLCNV.....VDLEKDCRRLRDPFPCF 383

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool/US10661430/runat_06012006_123135_6248/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10661430@CGN_1_1_5315@runat_06012006_123135_6248 -NCPU=6 -ICPU=3
-NO WMAP -LANG=QUEY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_hic:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_esc7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1212.5	59.7	860	7	CV125688
2	1148.5	56.5	866	7	CV125689
3	727.5	35.8	1590	4	CR614082
4	717	35.3	1196	11	DD035706
5	717	35.3	1860	4	CR612341
6	717	35.3	1870	4	CR607487
7	717	35.3	1876	4	CR592876

8	717	35.3	1901	4	CR599499	full-length
9	717	35.3	1979	4	CR925982	Pongo pyg
10	693.5	34.1	911	7	CR444324	
11	692.5	34.1	1738	4	BC004095	Mus muscu
12	680.5	33.5	970	1	AL558425	
13	675	33.2	791	1	AL558425	
14	675	33.2	803	8	CR744001	JGI ANBT1
15	670.5	33.0	796	8	CR757659	AGENCOURT
16	670.5	33.0	955	5	BO957788	AGENCOURT
17	669.5	33.0	879	6	CD514077	AGENCOURT
18	667.5	32.9	2369	6	BC029303	Mus muscu
19	664	32.7	737	8	CR744145	JGI ANBT1
20	658.5	32.4	844	1	AL866706	AL866706
21	658	32.4	687	5	BM054310	BM054310
22	657	32.3	759	8	CA457513	JGI XZG54
23	655.5	32.3	922	6	CP239240	AGENCOURT
24	651	32.1	907	5	BX848994	EX848994
25	649	32.0	893	8	CA362564	JGI XZ142
26	648.5	31.9	871	7	CK188001	EST77316
27	647	31.9	803	8	CR780638	AGENCOURT
28	646	31.8	1700	4	CR607401	full-length
29	645.5	31.8	876	5	BO929587	AGENCOURT
30	642.5	31.6	2303	4	CR749555	Homo sapi
31	638	31.4	661	3	BU617372	BU617372
32	638	31.4	1136	1	AL556185	AL556185
33	636.5	31.3	1035	10	AY410571	AY410571
34	634.5	31.2	857	8	CA442498	JGI XZG61
35	634	31.2	684	5	BM437339	BM437339
36	633	31.2	1368	4	CNS06MSO	Tetradon
37	633	31.2	2094	4	AK085368	Mus muscu
38	633	31.2	2103	4	AK052013	Mus muscu
39	633	31.2	2104	4	AK041692	Mus muscu
40	633	31.2	2104	4	AK050005	Mus muscu
41	633	31.2	2169	4	AK053603	Mus muscu
42	632.5	31.1	907	6	CF552958	AGENCOURT
43	628	30.9	622	5	BM447970	BM447970
44	623.5	30.7	1027	10	AY410573	Mus muscu
45	623	30.7	679	3	BU061791	BU061791

ALIGNMENTS

RESULT 1
LOCUS CV125688/c
DEFINITION OSTR30043C08 PDONR201 Entry vector Caenorhabditis elegans cDNA,
mRNA sequence.
ACCESSION CV125688
VERSION CV125688.1 GI:51716887
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 860)
Lamesch,P., Malsrein,S., Hao,T., Rosenberg,J., Li,N., Segueria,R.,
Boesak,S., Doucette-Stamm,L., Vandenhaute,J., Hill,D.E. and Vidal,M.
C. elegans ORFome Version 3.1: Increasing the coverage of ORFome
resources with improved gene predictions
Genome Res. (ORFome issue) (2004) In press
CONTACT: Philippe Lamesch and Tong Hao
Marc Vidal Lab
DPCI
44, Binney Street, Boston, MA 02115, USA
Tel: 6176323910
Fax: 6176325739
Email: philippe_lamesch@dfci.harvard.edu
PCR Primers
FORWARD: TGGCTTTCGTCATTGGCAGT
BACKWARD: TAAACAACGTTGAAGTCCCT.
Location/Qualifiers
1..860
/organism="Caenorhabditis elegans"

```

/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="male and hermaphrodite"
/tissue_type="all tissues"
/cell_type="all cells"
/dev_stage="all stages"
/clone_lib="pDONR201 Entry vector"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,49e-135 Length: 860
Score: 1212.50 Matches: 248
Best Local Similarity: 93.23% Conservative: 0
Best Local Similarity: 93.23% Mismatches: 9
Query Match: 59.70% Indels: 9
DB: 7 Gaps: 3

US-10-661-430-1 (1-383) x CV125688 (1-860)

QY 127 GULYSELITYRPRASP---THRIEALA---GLYGLYHIGLY-METPROLYSAPCY 144
DB 788 GAAATTCATATCCGAGCACCCGCCGCCGAGTACATGAAATGCCCTTAAGGAT 729
QY 144 s---ValAlaArg-HisArgValAlaIle-IleVal-ProTYrArgAsp-ArgGluAlaH 162
DB 728 TGGTGTTCAGAGGCGATCGTGTTCATATTATGTGCCCCCTATAGAGATTCGTAAGCAC 669
QY 162 ILEUVALGILEMELUHS-AsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTYR 181
DB 668 ATTGGAAATTAAGCTCCAAATAATTGCATCGTGTCCCAACAAACAAATGGACTAT 609
QY 182 AlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsn 201
DB 608 GCATTTTCATGTGGAGCAAGTGCGAATCAAGCTTATATCGCGGAAACTAATGAAAC 549
QY 202 ValGlyTyraAspValAlaSerArgLeuTYR-ProTYrGlnCysPheIlePheHisAspVal 221
DB 548 GTTGGATACAGACGTAGATCAAGCTCTACCAATGGCAGGCTTCATCTTCATGATGTC 489
QY 222 AspLeuLeuProGluAspAspArgAsnLeuTYRThrCysProIleGlnProArgHisMet 241
DB 488 GATTTCGCCCAAGATGACCGTAACCTGTACACGCTCAATTCAAACACGTCATATG 429
QY 242 SerValAlaIleAspLysPheAsnTYRLeuLeuProTYrSerIleIlePheGlyGlyIle 261
DB 428 AGGTATCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 369
QY 262 SerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTYrGly 281
DB 368 AGTGCACTAACAAAGATCACTGTAAAGAAATCAATGATTTTTCGATGATTTTGGGGT 309
QY 282 TTTGTYGLYGLUAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSer 301
DB 308 TGGGGCGGAGAGGCGACGATTTGGCGACGAAACATGATGCTGACCTGAAGAATTTCA 249
QY 302 ArgTYRProThrGlnIleAlaArgTYRLeuMetIleLysHisSerThrGlnAlaThrAsn 321
DB 248 AGATATCCGACAAATTCGACGATTAATAATTAATTAATTAATTAATTAATTAATTAAT 189
QY 322 ProValAsnLysCysArgTYRLeuIleMetGlyGlnThrLysArgArgCTTTPThrArgAsp 341
DB 188 CCAATTAAATATCCGCTACAAATAATATGGCCAAAGCAAGCCGATGGAACGCGAC 129
QY 342 GlyLeuSerAsnLeuLysTYRLeuValAsnLeuGlnLeuLysProLeuTYRThrArg 361
DB 128 GGCCTTAAGCAATCGAAGTATTAAGCTCGTAAATTCGAATTTGAAGCCTCTCTACACTGA 69
QY 362 AlaValAlaAspLeuGlnLysAspCysArgArgGlnLeuAlaGlyAspAspPheProThr 381
DB 68 GCGCTCGTCATTTGCTGAAAAAGACTGCGCGGAGAGCTGCGAAGGACTTTTCAACG 9
QY 382 CysPhe 383

```

```

Db 8 TGTATT 3
RESULT 2
CV125689
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Marc Vidal Lab
DFCI
44, Binney Street, Boston, MA 02115, USA
Tel: 6176323910
Fax: 6176325739
Email: philippe_james@dfci.harvard.edu
PCR Primers
FORWARD: TGGCTTTTCGTCATTTGGCAGT
REVERSE: TAAACACGTTGGAAGTCCCT.
Location/Qualifiers
1..866
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/sex="male and hermaphrodite"
/tissue_type="all tissues"
/cell_type="all cells"
/dev_stage="all stages"
/clone_lib="pDONR201 Entry vector"

ORIGIN
Alignment Scores:
Pred. No.: 1.39e-127 Length: 866
Score: 1148.50 Matches: 236
Percent Similarity: 92.08% Conservative: 8
Best Local Similarity: 89.06% Mismatches: 14
Query Match: 56.55% Indels: 7
DB: 7 Gaps: 3

US-10-661-430-1 (1-383) x CV125689 (1-866)

QY 2 AlaPheArgHisLeuAlaValAlaArgLysSerLeuLeuValLeuCysAlaValLeu 21
DB 3 GCTTTTCGTCATTTGGCAGTCGCAACCAATCAAGTCTGTGCTGCTACTTTGTGCGCTTCT 62
QY 22 LeuLeuValHisAlaMetIleTYRLeuLysIleProSerLeuTYRLeuLeuLeuThrIleGly 41
DB 63 CTATTGTTCATGCAATGATTTTAATAATTCATCGCTTTACAAAGAACTTACTATTCGGC 122
QY 42 SerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSer 61
DB 123 TCTTCGACCTTATTTGCGAGTCGTAACGCAATGAAGGACGTGCTCGGAAATACGGCTTCC 182
QY 62 ThrSerAspAspLeuLeuAspThrTYrAsnSerThrPheSerProIleSerGluValAsn 81
DB 183 ACTTCGAGATATTACTTGATGATGTAAGTGAATTCACAGTTTTCACGATTTCTGAAGTTAAT 242
QY 82 GlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGln 101
DB 243 CAAACTAGTTTATGAGGACATTCGTCCAATCTGTTCCTCCGCAACACGACTTTTCAA 302

```

```

QY 102 PheCysAenGIInThrProHishLeuValGlyProIleArgValPheLeuAspGluPro 121
| | | | |
DB 303 TTCTGTATACAGACACCTCCCACTCGTCGAGACCATCGGTATTTCTTCATGAGCC 362
| | | | |
QY 122 AspPheLeuThrLeuGluValIleTyrProAspThrHisAlaGlyHisGlyMetPro 141
| | | | |
DB 363 GACTTCAAAACCTCGAGAAAATCTATCCGACACGACGCGCGGTGACATGGAATGCT 422
| | | | |
QY 142 LysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAla 161
| | | | |
DB 423 AAGGATAGTGTGCAAGGATCGTGTGCTATTAATGTCCTCATAGAAATGTAAGCA 482
| | | | |
QY 162 HisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnIleLeuAspTyr 181
| | | | |
DB 483 CATTTGAGAAATATATGCTCCACATTTGCACTGCTTCTGCCAACAACATTTGACAT 542
| | | | |
QY 182 AlaIlePheIleValGluGlnValAlaAenGIInThrPheAsnArgGlyLysLeuMetAsn 201
| | | | |
DB 543 GCAATTTTCATTGTGAGCAAGTGGGATCAGACGTTAATCGCGGAAAATGAATGAA 602
| | | | |
QY 202 ValGlyTyrAspValAlaSerArgLeuTyrProTyrGln-CysPheIlePheHisAspVal 221
| | | | |
DB 603 GTTGGATAGAGATGATCCGCTCTACCCATGAGGAGGCTTCTCTTCATGATGT 662
| | | | |
QY 221 L-AspLeuLeuProGlu-AspAspArgAsnLeuTyr---ThrCysProIleGlnPro-Arg 239
| | | | |
DB 663 CCGATTTTACTGCCCGAAAATGACCTGATACCTGGTACCCGTGTCATTTAACAACACCG 722
| | | | |
QY 239 GHisMetSerValAla---IleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePhe 258
| | | | |
DB 723 TCATATGAGGTGTAGCCGATCCGAAAATTTCAATTAAAAACCT---TCCCTTATTTC 779
| | | | |
QY 258 egllygl 260
| | | | |
DB 780 GGGGGGA 786
| | | | |

```

```

RESULT 3
CR614082 1590 bp mRNA linear HTC 21-JUL-2004
LOCUS Full-length cDNA clone CS0D005YK01 of T cells (Jurkat cell line)
DEFINITION Cot 10-normalized of Homo sapiens (human).
ACCESSION CR614082
VERSION CR614082.1 GI:50494889
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1590)
Li W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Paradey Avenue
2 (bases 1 to 1590)
Genoscope.
REFERENCE
AUTHORS Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life Technologies, a
division of Invitrogen.
FEATURES
source
1..1590
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/cloname="CS0D005YK01"
/tissue_type="T cells (Jurkat cell line) Cot
10-normalized"
/plasmid="pCMVSPORT_6"

```

ALIGNMENT Scores:

```

Pred. No.: 2.98e-76 Length: 1590
Score: 727.50 Matches: 142
Percent Similarity: 62.21% Conservative: 49
Best Local Similarity: 46.25% Mismatches: 112
Query Match: 35.82% Indels: 4
DB: 4 Gaps: 3

```

US-10-661-430-1 (1-383) x CR614082 (1-1590)

```

QY 64 AspAspLeuLeuAspThrThrAsnSerThrPheSerProIleSerGluVal---AenGIIn 82
| | | | |
DB 44 GATTCGCTGCGAGTGTGATGAGAGCCTTCTCTGCTCTGTCATTTCCAAATTTT 103
| | | | |
QY 83 ThrSerPheMetGluAspIleArgProIleLeuPhe-ProAspAsnGIInThrLeuGlnPh 102
| | | | |
DB 104 AGATTAATGCTCACATCTCTGTCCCGAGGGGCTCTCAAGCTCTCAAGGTCCCTA 163
| | | | |
QY 102 eCysAenGIInThrProHishLeuValGlyProIleArgVal---PheLeuAspGluPr 121
| | | | |
DB 164 CTGTCCAGAAAGATCTCTCTCTTAATGTTGGTCTGTGTGGGTCTTTAGCCAGTGCC 223
| | | | |
QY 121 AspPheLeuThrLeuGluValIleTyr-ProAspThrHisAlaGlyHisGlyMetPr 141
| | | | |
DB 224 ATCATGTCAGAGATTTGTGAGCGGAAATCCCGGCTAGAACCGGGGCGGTACCGCC 283
| | | | |
QY 141 LysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAla 161
| | | | |
DB 284 TCACGTTGTAGACCCCGCTCCGAAACAGCAATCATATGCTCATGTCGTCGCGGAGCA 343
| | | | |
QY 161 HisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnIleLeuAspTyr 181
| | | | |
DB 344 CCACCTGGCGCTCGCTCTCAACCTGCAACCCCTTGTGACGGCCAGCGATGCTGCTTA 403
| | | | |
QY 181 rAlaIlePheIleValGluGlnValAlaAenGIInThrPheAsnArgGlyLysLeuMetAs 201
| | | | |
DB 404 TGGCATATGATCATCAACCAACGAGCTGAAATGAAATTTAAACAGGGCAAACTGTGAA 463
| | | | |
QY 201 nValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVal 221
| | | | |
DB 464 GCTTGAGGTCGAGAGGCTCTGCGTGAATGAAGATGGAGCTGCTTCTTGACAGATGT 523
| | | | |
QY 221 AspLeuLeuProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHis 240
| | | | |
DB 524 GGAACCTTTGCCAAGAAAATGACACAACTGTATGTGTGTGAACCCCGGGAGACCCGCCA 583
| | | | |
QY 240 sMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyGI 260
| | | | |
DB 584 TGTTCCTGCTGTATGAACAAGTTGATACAGCTCCCGTACCCCGAGTACTTGGAGAG 643
| | | | |
QY 260 YIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGIInPheSerAsnAspPhnTr 280
| | | | |
DB 644 AGTCTACACACTTACTCTCTGACACAGTACCTGAAGATGAATGAGCTTCCCAATGAATACG 703
| | | | |
QY 280 pGlyTyrpGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysVala 300
| | | | |
DB 704 GGGCTGGGGGTGTAGATGATGACGACATTTGACAGAGGTGGCCCTGGGTGATGAAT 763
| | | | |
QY 300 lSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaTh 320
| | | | |
DB 764 CTCTGCGCCGCCACATCTGTAGACACTATTAAGTGTGAAGCACCGAGAGATAGAGG 823
| | | | |
QY 320 rAsnProValAsnLysCysArgTyrLysIleMetClyGlnPhLysArgArgTyrThrAr 340
| | | | |
DB 824 CAATAGAGAAATCCCAACAGATTTGACCTCTGCTGTCGTCACCAATTCCTGGACGGA 883
| | | | |
QY 340 GAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrTh 360
| | | | |

```


Db	984	AGATGGGATGAACCTACTGACATCACTGACTTCTGCTGCAAGAGCTGGGGCTCTTTATAC	943
Qy	360	TARGALaValValAepIieu 366	
Db	944	CAACATCAGACGACAGACTT 962	
RESULT 4			
LOCUS	DQ035706		
DEFINITION	DQ035706 Homo sapiens B4GALT3 gene, VIRUTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	DQ035706		
VERSION	DQ035706.1		
KEYWORDS	GI:66866915		
SOURCE	GSS.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 1196) Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civejlo,D., White,T.J., Sinsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.		
FEATURES	Location/Qualifiers		
SOURCE	1..1196 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="1" /gene="B4GALT3" /locus_tag="HC4901"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,61e-75	Length:	1196
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	11	Gaps:	2
US-10-661-430-1 (1-383) x DQ035706 (1-1196)			
Qy	95	ProAARaNGlThrluGlnPhcCySaNGlInThrProProHlenuValAlProIle	114
Db	219	CCAGCTCTCAAGTGTGCCCTACTGTCCAGAACATCTCTCTTACTGGAGTCTGG	278
Qy	115	ArGVal----PheIuAaRgIuProAaRPhelYsThrluGluIuYrPrAaRPhr	133
Db	279	TGGGTGTCTTTAAGCCCACTGCAATCACTGCGAGAAATGTGGAGCGAAATCCCGGGA	338
Qy	134	HisAlaGluYlYHlGluYMetProIuYaARCySaValAlaYgHlAaYValAlaIleIle	153
Db	339	GAACCAAGGGGGCGGATACCGCCCTGCAAGGTGTGAAGCCCGCTCCGAACAGCCATCAT	398
Qy	154	ValProTYArGdaARaRgIuAlaHlGluAaYglIeMetLeuHlAaYHlGluHlGluSerIeu	173

Df		GTGGCTCATGTCGCCGGAGACACCACTGGCGCTGCCTCACACTGCACCCCTTC	458
OY		174 LeuAlaIysgInGlnLeuAepTyrAlaIlePheIleValGlunIvalAlaAsnGlnThr	193
Df		459 TTGGACGGCCAGCAGCTTGCTTAAGGCATCTATTATCCACCAGCGCTGGAAATGAACA	518
OY		194 PheAsnArgGlyLysLeuMetAsnValIGlTYrAspValAlaSerArgLeuTyProTrp	213
Df		519 TTTAACAGGGCGAAACCTGTGAACGTTGGGGTGGAGAGGCCCTCGTAGTAAGAAGTGG	578
OY		214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThr	233
Df		579 GACTGCGCTGTTCTTGACAGATGTGACCTCTTCCAGAAAATGACCAACATCTGTATGTG	638
OY		234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTYrLysLeu	252
Df		639 TGTAACCCCCGGGAGACCCCGCAGATGGCCGTGTCTATGAACAAGTTGGATAAGCCTC	698
OY		253 ProTYrSerAlaIlePheGlyGlyIleSerAlaleuthrLysAspHisLsleuLYsIle	272
Df		699 CCGRACCCCACAGTACTTGGAGAGGTCTACACACTTACTCTCTGCACGATCCTGAAGATG	758
OY		273 AsnGlyPheSerAsnAspPheTrpGlyTYrGlyGlyLubAspAspAspLeuAlaITHArg	292
Df		759 AATGGCTTCCCCANATGAATACTGGGGGTGGGTGTGTAGATGACGACATGTGTACAGG	818
OY		293 ThrSerMetAlaGlyLeuLysValSerArgTYrProThrglnIleAlaArgTYrLysMet	312
Df		819 GTGGCGCTCGTGGGAGATGAGATCTCTCGCCCCCACATCTGTAGGACATTAAGAATG	878
OY		313 IleLysHisSerThrGluAlaTHAsnProValAsnLysCySArgTYrLysIleMetGly	332
Df		879 GTGAAGCACCGAGAGATAGGCGCAATGAGAAATCCCCACAGATTTGACCTCTGTCTC	938
OY		333 GlnThrLysArgArgTYrThrArgAspGlyLeuSerAsnLeuLysTYrLysLeuValAsn	352
Df		939 CGTACCCAGAAATCTCTGGACGCAAGATGGATGAACATCATGACATACCAAGTGTGGCT	998
OY		353 LeuGlnLeuLysProLeuTYrThrArgAlaValValAspLeu 366	
Df		999 CGAGAGCTGGGGCCTCTTATATACCAATCACAGACAGACATT 1040	
RESUT 5			
Cd	CR612341	1860 bp mRNA linear HTC 21-JUL-2004	
LOCUS	full-length cDNA clone CSDDK011M18 of Hela cells Cot 25-normalized		
DEFINITION	of Homo sapiens (human).		
ACCESSION	CR612341		
VERSION	CR612341.1 GI:50493148		
KEYWORDS	HTC; CNSLT_cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homniidae; Homo		
REFERENCE	1 (baaes 1 to 1860)		
AUTHORS	Li,W.B., Gruber,C., Jeessee,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
REMARK	Contact : Feng Liang Email : fliang@lifeitech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue 2 (baaes 1 to 1860) Genoscope. Direct Submission Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr - Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oIigo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and clone into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		
COMMENT			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

FEATURES

Location/Qualifiers
1. .1860
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YH18"
/tissue_type="Hela cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 7.04e-75 Length: 1860
Score: 717.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.30% Indels: 2
DB: 4 Gaps: 2

US-10-661-430-1 (1-383) x CR612341 (1-1860)

```
OY 95 ProAaPaaEngInThrLeuGlnPheCysAaEngInThrProH1sLeuValGlyProIle 114
Db 421 CCAGCTCTCAAGAGTCTGCCCTTACCTCCAGAACATCTCTCTTATGAGGCTCTG 480
OY 115 ArgVal---PheLeuAaPgluPaaPaaPheLysThrLeuGlnLysIleTyPProAaPThr 133
Db 481 TCGGTGCTCTTAAAGCCAGTGCATCTGAGAGATTTGTGAGCGGAATCCCGGGA 540
OY 134 HisAlaGlyGlnHisGlyMetProLysAaPValAlaArgHisArgValAlaIleIle 153
Db 541 GAACCAAGGGGCGCGGTACCCGCTGACAGGTTGTAGCCCGCTCCGGAACAGCCATCAT 600
OY 154 ValProTyArgAaPaaPaaGlnAlaHisLeuAaGlyIleMetLeuHisAaLeuHisSerLeu 173
Db 601 GTGCTCTCAAGTCCCGGAGAGACCACTGCGCTGCTGCTACCACTGACCTGACCTTC 660
OY 174 LeuAlaLysGlnGlnLeuAaPtyrAlaIlePheIleValGlnValAlaAaGlnInThr 193
Db 661 TTGACAGCGCAGACACTTCTTATGGCATCTATGTCATCCACAGCGCTGGAATGGAACA 720
OY 194 PheAaAaGlyGlyLysLeuMetAaAaValGlyTyPaaPaaAlaSerArgLeuTyPProTyr 213
Db 721 TTAAACAGGGCAAACTGTGAACGTTGGGTCGAGAGCCCTGCTGATGAGAAGTGG 780
OY 214 GlnCysPheIlePheHisAaPValAaPLeuAaPProGlnAaPaaPaaPaaPaaPaaPaa 233
Db 781 GACTGCTGCTTCTTGCATGATGAGCTCTTGCAGAAATGACCAATATGTATGTG 840
OY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAaPtyPheAaAaTyPLeu 252
Db 841 TGTGACCCCGGGGAGACCCGCGCATGTGCGCTTGTCTATGACAAAGTTTGGATTCACCTTC 900
OY 253 ProTySerAlaIlePheGlyGlyIleSerAlaLeuThrLysAaPHisLeuLysIle 272
Db 901 CCGTACCCCGCAAGTACTGAGAGATCTGACACTTACTCTGACAGTACCTGAAAGATG 960
OY 273 AaAaPtyPheSerAaPaaPaaPheTyrGlyTyrGlyGlyValAaPaaPaaPaaPaaPaaPaa 292
Db 961 AATGCTCTTCCCAAGAACTATGAGGCTGAGGAGTGTGAGAGATGACACATCTCTACAGG 1020
OY 293 ThrSerMetAlaGlyLeuLysValSerArgTyPProThGlnIleAlaArgTyPLeuMet 312
Db 1021 GTGGCGCTGAGTGGATGAAATCTCTGCGCCCGCCACATCTGTGACACATTAAGATG 1080
OY 313 IleLysHisSerThGlnAlaThraAaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 332
Db 1081 GTGAAGACCCGAGAGATTAAGGACATGAGAAATCCCGACAGATTTGACCTCTGATC 1140
OY 333 GlnThrLysAaArgTyPThraAaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 352
Db 1141 CGTACCCAGAAATCTCTGAGCAGAAATGAGATGAATCTCACTACATCAAGTTGCTGCT 1200
OY 353 LeuGlnLeuLysProLeuTyPThraGlnAlaValAaPLeu 366
```

Db 1201 CGAGAGCTGGGGCTCTTTATATCCAAATCACAGACAGAT 1242

RESULT 6

LOCUS

DEFINITION

Full-length cDNA clone CS0DC015YB21 of Neuroblastoma Cot 25-normalized of Homo sapiens (human).

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

QY 194 PheAsnArgLylyseuMetAsnValGlyTyAspValAlaSerArgLeuTyProThr 213
DB 733 TTTAAAGAGGCAAAACTGTTGAACGTTGGGTCGAGAGGCCCTCGCTGATGAAGAGTGG 792
QY 214 GlnCysPheIlePheHisAspValAspLeuProGluAspAspArgAsnLeuTyThr 233
DB 793 GACTGCTGCTTTCGACGATGTGAGACTCTTCCAGAAAAATGACCAACATCTGTATGTG 852
QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspIlyPheAsnTyIlyIleu 252
DB 853 TGTGACCCCGGGAGACCCCGCCATGTTCCGTTGATGAAGCAAGTTGGATCAGCCCTC 912
QY 253 ProTySerAlaIlePheGlyIlyIleSerAlaIleThrIlyAspHisIleuIlyIle 272
DB 913 CCGTACCCCAAGTACTTCGAGAGAGTCTGACACTTACTCTTCAACAGTACCTGAAGATG 972
QY 273 AsnGlyPheSerAspAspPheTyTPGlyTyIlyGlyIlyAspAspAspLeuAlaThrArg 292
DB 973 AATGGCTTCCCAATGAAATCTGGGGCTGGGGTGGTGAAGATGACGACATGTCTACAGG 1032
QY 293 ThrSerMetAlaGlyIlyleuIlyValSerArgTyProThrGlnIleAlaArgTyIlyMet 312
DB 1033 GTGCGGCTGCTGGATGAAGATCTCTCGGCCCCCACTGTGTAGACACTTAAGATG 1092
QY 313 IleIlyHisSerThrGlnAlaThrAsnProValAsnIlyCysArgTyIlyIleMetGly 332
DB 1093 GTAAAGACCCGAGAGATTAAGGCAATAGAGAAATCCCAAGATTTCACCTCTGCTG 1152
QY 333 GlnThrIlyAspArgTyTPGlyTyAspArgIlyleuSerAsnLeuTyIlyIlyIleuValAsn 352
DB 1153 CGTACCAGAAATCTCTGACGCAAGATGGATGAATCACTGACATCTCAGTGTGCTGGCT 1212
QY 353 LeuGlnLeuIlyProLeuTyThrThrArgAlaValAlaAspLeu 366
DB 1213 CGAGAGCTGGGGCTTTTATACCAACATCAGACAGACAGACTT 1254

RESULT 7
CR592876
LOCUS 1876 bp mRNA linear HTC 21-JUN-2004
DEFINITION full-length cDNA clone CSOD1008YC21 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR592876
VERSION CR592876.1 GI:50473683
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1876)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue
2 (bases 1 to 1876)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
location/qualifiers
1..1876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source

Alignment Scores:
Pred. No.: 7.13e-75 Length: 1876
Score: 717.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.30% Indels: 2
DB: 4 Gaps: 2

US-10-661-430-1 (1-383) x CR592876 (1-1876)

ORIGIN
/clone="CSOD1008YC21"
/cisse_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

QY 95 ProAspAsnGlnThrIleuGlnPheCysAsnGlnThrProProHisIleuValGlyProIle 114
DB 434 CCAAGCTCTCAAGAGTCTGCCCTACTGTCGAAACGATCTCTCTTAACTGGAGTCTGTG 493
QY 115 ArgVal---PheIleuAspGluProAspPheIlyThrIleuGlnIlyIleTyProAspThr 133
DB 494 TCGGTCTCTTTAGCCCACTGCTCATCTGCGAGAGATGTGGAGCGGAATCCCGGGA 553
QY 134 HisAlaGlyIlyHisGlyMetProIlyAspCysValAlaArgHisIlyArgValAlaIleIle 153
DB 554 GAACAGGGGGCCGGTACCGCCCTGCGAGGTGTGAAGCCCGCTCCCGAAGCCATCATTT 613
QY 154 ValProTyArgAspArgGlnAlaHisIleuArgIlyIleMetLeuHisAsnIlyHisSerLeu 173
DB 614 GTGCTCTATGTGTCCCGGAGACACACTGCGCTGCTCTCAACCTGACCGCCCTTC 673
QY 174 LeuAlaIlyGlnIleuAspTyAlaIlePheIleValGlnIleValAlaAsnGlnThr 193
DB 674 TTGCAGCGCCAGAGAGTCTTATGSCATCTATGTCATCCACAGGCTGGAATGACACA 733
QY 194 PheAsnArgLylyseuMetAsnValGlyTyAspValAlaSerArgLeuTyProThr 213
DB 734 TTTAAAGAGGCAAAACTGTTGAACGTTGGGTCGAGAGGCCCTCGCTGATGAAGAGTGG 793
QY 214 GlnCysPheIlePheHisAspValAspLeuProGluAspAspArgAsnLeuTyThr 233
DB 794 GACTGCTGCTTTCGACGATGTGAGACTCTTCCAGAAAAATGACCAACATCTGTATGTG 853
QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspIlyPheAsnTyIlyIleu 252
DB 854 TGTGACCCCGGGAGACCCCGCCATGTGGCGTTGCTATGAACAAGTTTGGATACAGCCTTC 913
QY 253 ProTySerAlaIlePheGlyIlyIleSerAlaIleuThrIlyAspHisIleuIlyIle 272
DB 914 CCGTACCCCAAGTACTTCGAGAGAGTCTCAGCACTTACTCTGACCAAGTACCTGAAGATG 973
QY 273 AsnGlyPheSerAspAspPheTyTPGlyTyIlyGlyIlyAspAspAspLeuAlaThrArg 292
DB 974 AATGGCTTCCCAATGAAATCTGGGGCTGGGGTGGTGAAGATGACGACATGTCTACAGG 1033
QY 293 ThrSerMetAlaGlyIlyleuIlyValSerArgTyProThrGlnIleAlaArgTyIlyMet 312
DB 1034 GTGCGGCTGCTGGATGAAGATCTCTCGGCCCCCACTGTGTAGACACTTAAGATG 1093
QY 313 IleIlyHisSerThrGlnAlaThrAsnProValAsnIlyCysArgTyIlyIleMetGly 332
DB 1094 GTGAACACACGAGAGATTAAGGCAATGAGAAAAATCCCAAGATTTCACCTCTGCTG 1153
QY 333 GlnThrIlyAspArgTyTPGlyTyAspArgIlyleuSerAsnLeuTyIlyIlyIleuValAsn 352
DB 1154 CGTACCAGAAATCTCTGACGCAAGATGGATGAATCACTGACATCTCAGTGTGCTGGCT 1213
QY 353 LeuGlnLeuIlyProLeuTyThrThrArgAlaValAlaAspLeu 366
DB 1214 CGAGAGCTGGGGCTTTTATACCAACATCAGACAGACAGACTT 1255

RESULT 8
CR599499

gene 3 (Homo sapiens) "
1. .1979
/gene="DKFZp459J2220"

CDS

271. .1452
/gene="DKFZp459J2220"

/codon_start=1
/product="hypothetical protein"

/protein_id="CA129630.1"

/db_xref="GI:56403664"

/translation="MLRRLLEPCTALLVLSQALVMMYLSLGGPRLSALRGGDGP
TFDYSHPRDVTSLSHPGAFGGPPAPQGLPYCPERSPLLVGVSVSFSVPLAIV
ERNRVRVPGGRYPACGPRSRITAIIVPHARHHLRLYLHLAPLQROQLAVGILYV
HQAGNGTFNRALINLVGREALRDEMDCLPLADVDLLEPNDNLVYCDPRGRHVA
VANKKGYSLPYPOYFGVYALTPDDYLLKNGGPNEXYMGWGGEDDILATRVLAGMKI
SRPSTGYKGVKVGHRGDKGENENPHRFDLLVTRTNSMTQDGNLSLYOLLAEGLP
YTNITADIGTRPRAPRPSGRYPSPSSQAFRQEMLRPRAPRPPRANHTALRGS
H"

ORIGIN

Alignment Scores:

Pred. No.:	7,73e-75	Length:	1979
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	4	Gaps:	2

US-10-661-430-1 (1-383) x CR925982 (1-1979)

```

QY 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIle 114
Db 475 CCAGCTCCTCAAGGTCGCCCTACTGTCCAGAACATCTCTCTTAAGTGGGCTGTG 534
QY 115 ArgVal---PheLeuAspGluProAspPheLeuThrLeuGluValIleTyrrProAspThr 133
Db 535 TCGGTGCTCTTACGCCAGTGCCTCACTGGCAGAGATGTGTGAGCGGAAATCCCGGGA 594
QY 134 HisAlaGlyLysIleGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 595 GAACACGGGGGCGCGTACCGCCCTGCAGGTTGTAGCCGCCCTCCGACACGCATCAT 654
QY 154 ValProTyrrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 655 GTGCCTCATGTGCGCCGGAGACCACTGCGCTGTACCTGACCTGACCCCTTC 714
QY 174 LeuAlaGlyGlnGlnLeuAspTyrrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 715 TTGCAGCGCCAGCGCTTGTATGCACTTATGTCAATCCACGAGCTGGAATGACACA 774
QY 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrrAspValAlaSerArgLeuTyrrProTyr 213
Db 775 TTTRACAGCGCAAACTGTGAACCTTGGGGTGGAGAGCGCCGCGTGAATGAAGATGG 834
QY 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrrThr 233
Db 835 GATGCTGCTTCTTGAATGTGATGTGACCTTGTGCACAAATAATGACCAATCTGTATGTG 894
QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrrLeu 252
Db 895 TGTACACCCCGGAGACCCCGCATGTGCGGTCTGTGAACAAGTTTGATACAGCCTC 954
QY 253 ProTyrSerAlaIlePheGlyLysIleSerAlaLeuThrLysAspHisLeuLysIle 272
Db 955 CCGTACCCCGCATCTTCCGAGAGAGTCTCGCGCTTACTCTGACACAGTACGTGAAGATG 1014
QY 273 AsnGlyPheSerAsnAspPheTyrrGlyTyrrGlyLysAspAspAspLeuAlaThrArg 292
Db 1015 AATGCTTCCCATGATGATCTGGGGCTGGGGTGTGATGATGATGATGATGATGATGATGATG 1074
QY 293 ThrSerMetAlaGlyLeuValSerArgTyrrProThrGlnIleAlaArgTyrrLysMet 312
Db 1075 GTGCGCTGCTGGAGATGATCTTCGCGCCCGCCACATCTGTAGACACTATTAAGATG 1134

```

QY 313 IleLysHisSerThrGlnAlaThrAsnProValAsnLysCysArgTyrrLysIleMetGly 332

Db 1135 GTGAAGCACGAGAGATTAAGGGCAACGAAATATCCCAAGATTTAGCTCTGTC 1194

QY 333 GlnThrLysArgArgTyrrThrArgAspGlyLeuSerAsnLeuLysTyrrLysLeuValAsn 352

Db 1195 CGTACCCCAAAATTCCTCGAGCAGAAAGTGGATGATCACTACATACCAATTTGCTGGCT 1254

QY 353 LeuGluLeuLysProLeuTyrrThrArgAlaValAlaAspLeu 366

Db 1255 CGAAGATTGGGCTCTTTATACCAATCATCAGCAGACAT 1296

RESULT 10

CR444324 911 bp mRNA linear EST 19-JUN-2004

LOCUS CR444324

DEFINITION CR444324 XGC-tailbud Xenopus tropicalis cDNA clone TTBA074c18 5',

MRNA Sequence.

CR444324

CR444324.1 GI:48969911

EST.

Xenopus tropicalis (western clawed frog)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

Xenopodinae; Xenopus; Silurana.

REFERENCE

1 (bases 1 to 911)

Croning,M.D.R., Aehurst,J.L., Taylor,R., Garrett,N. and Rogers,J.

Unpublished (2004)

CONTACT: Croning MDR

Address: Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: TTBA074c18.plkS6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Nigel Garrett.

Seq primer: Sp6.

Location/Qualifiers

1..911

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8164"

/clone="TTBA074c18"

/dev_stage="tailbud (stage 28-30)"

/lab_host="Escherichia coli DH10B."

/clone_lib="XGC-tailbud"

/note="Vector: pCS107; Site_1: BcoRI; Site_2: NotI; cDNA

was oligo dt primed from 5ug of poly A+ RNA from tailbud.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Alignment Scores:

Pred. No.:	1.66e-72	Length:	911
Score:	693.50	Matches:	125
Percent Similarity:	65.95%	Conservative:	59
Best Local Similarity:	44.80%	Mismatches:	84
Query Match:	34.15%	Indels:	11
DB:	7	Gaps:	2

US-10-661-430-1 (1-383) x CR444324 (1-911)

```

QY 74 PheSerProIleSerGluValaGlnThrSerPheMetGluAspIleArgProIleLeu 93
Db 104 TAGCCAACTCAAGCTTAATCAAGTCTTTCCTTACGGGAA----- 145
QY 94 PheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyPro 113
Db 146 -----GAGAGTTGGCATTTCTGCGGAGACCTCACCTTATATAGGGGGCCA 193
QY 114 IleArgValPheLeuAspGluPro---AspPheLysThrLeuGluLysIleTyrrProAsp 132

```

```

Db      194 ATAGAGTCACGCTTCCGAGAGAACTTGTACTGAGAGAGTGGAGAAAAAGAACCTTAC 253
Qy      133 ThrHisAlaGlyGlyHisGlyMetProlyAspGlyValAlaArgHisArgValAlaIle 152
Db      254 GTGTCCAAAGAGAGGGGCTACCAAGCCAGCTGAGTCCACTATATAAAGCGGCGTG 313
Qy      153 IleValProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSer 172
Db      314 ATCATCCCTCACCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373
Qy      173 LeuLeuAlaLeuGlnGlnLeuAspTyrAlaAlaIlePheIleValGlnGlnValAlaAsnGln 192
Db      374 TTCTGACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
Qy      193 ThrPheAsnArgGlyTyrLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrPro 212
Db      434 ACTTTCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493
Qy      213 TTPGInCyPheIlePheHisAspValAlaPheLeuLeuProGluAspAspArgAsnLeuTyr 232
Db      494 TGGAGACTGTTGTTCTACACAGATGTCAGACCTCATTCCTCCGAGAGAGAGAGAGAGAGAT 553
Qy      233 ThrCyProIleGlnProArgHisMetSerValAlaAlaAspArgPheAsnTyrIleVal 252
Db      554 ACCTCGCAAAATTTCCCAACACCGCTGATCGCATGAGACAAAGTTTGATATCAAAATTG 613
Qy      253 ProTyrSerAlaIlePheGlyGlyIleSerAlaIleuThylAspHisIleuValIle 272
Db      614 CCTTCAAAATCTTACTTGGGAGAGTGCAGCCCTTCCCAAGAAATATCATGAAGATG 673
Qy      273 AenglyPheSerAspAspPheTTPGlyTTPGlyGlyIleAspAspAspLeuAlaThrArg 292
Db      674 AACGGTTCCTCCCAACAACTACTGAGATGGGCGGAGAGATGATATATAGCATCAAG 733
Qy      293 ThrSerMetAlaGlyLeuValSerArgTyrProThrGlnIleAlaArgTyrIleMet 312
Db      734 GTGCGCACTTACGGGAGCATGATCATCTCCGCTTATCATGACAGCGGAGATATAAATG 793
Qy      313 IleyHisSerThrGlnAlaThrAsnProValAsnLeuCyAspArgTyrIleMetGly 332
Db      794 ATCAAGACAGGCGCACCAAGGCGCAATGAGCAAAATCCCAAAAGTTTAAACATGCTGACG 853
Qy      333 GlnThrIleArgArgTyrThrArgAspGlyLeuSerAsnLeuIleTyrIleVal 351
Db      854 AAGACCCGCGGACGTCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 910

RESULT 11
BC004095      1738 bp      mRNA      linear      HTC 19-NOV-2003
LOCUS      Mus musculus UDP-gal-4-beta-galactosyltransferase,
DEFINITION      polypeptide 3, mRNA (cDNA clone IMAGE:3594143), containing
ACCESSION      BC004095
VERSION      BC004095.1 GI:14708685
KEYWORDS      HTC.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 1738)
Strusberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shennen,C.M., Schlier,G.D.,
Altschul,S.F., Zeeberg,B., Buecaw,K.H., Schaefer,C.F., Bhac,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachecho,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schertz,T.E., Brownstein,M.J., Usdin,T.B., Toshitaki,S.,
Carinci,P., Prange,C., Paha,S.S., Loggellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bonfard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Schnerich,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1738)
Strusberg,R.
Direct Submission
Submitted (28-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LINL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson,Mark) mcdopaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov
Series: IMAGE Plate: 11 Row: 1 Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES
source
1..1738
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3594143"
/tissue_type="mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_11b="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Alignment Scores:
Pred. No.: 5,85e-72 Length: 1738
Score: 692.50 Matches: 133
Percent Similarity: 64.60% Conservative: 44
Best Local Similarity: 48.54% Mismatches: 95
Query Match: 34.10% Indels: 3
DB: 4 Gaps: 2
US-10-661-430-1 (1-383) x BC004095 (1-1738)
Qy      95 ProAspArgGlnThrLeuGlnPheCyAsnGlnThrProGlnIleValGlyProIle 114
Db      264 CTTCCGGCTCAAGAGATGCTTACTGTGTCAGAAAGATACCTTTTACTGAGGCTCTG 323
Qy      115 ArgVal---PheLeuAspGlnProAspPheIleGlnGlyValIleTyrProAspThr 133
Db      324 TCAGTATCTTTAGCCCGGTCGTCACAGAGATGTCGAGATGTCGAGATCCCGGCG 383
Qy      134 HisAlaGlyGlyHisGlyMetProIleAspGlyValAlaArgHisArgValAlaIle 153
Db      384 GAATCAGGAGGCGGTACCTGTCAGAGGAGTGAAGCTCGCTCCGCAACCATTAATT 443

```



```

QY 154 ValProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
DB 444 GTGCCCCATGTCGCCGGAGGACACACTTCGCGCTGCTCATATCACTGCACCCCTTC 503
QY 174 LeuAlaLeuGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
DB 504 CTCACACGCGCAGAGCTTGCGTACGGCATTTATGTCATCCACACAGCGCTGGAAATGGAACG 563
QY 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyr 213
DB 564 TTTHACAGGCAAAAGCTGCTGAACGTAAGGCTGAGGAAAGCCCTTCGTATGTAAGATGG 623
QY 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr 233
DB 624 GACTGCTGTTGTTTACACGACGTCGACCTCTCCAGAAAACGACCTATACCTGATGTC 683
QY 234 Cys---ProIleGlnProAlaGlnIleMetSerValAlaIleAspLysPheAsnTyrLysLeu 252
DB 684 TGGACCCCGCGGAGACCCCGCCACGTTGCTGTGCGCATGAAACAAAGTTGA--TACAGCCTC 741
QY 253 ProTyrSerAlaIlePheGlnGlyIleSerValLeuThrLysAspHisLeuLysIle 272
DB 742 CCGTACCCCGCAGTACTTTGGCGAGTTTCAGCGCTCACTCCGACGCTGACCTGACATG 801
QY 273 AsnGlyPheSerAsnAspPheTyrGlyTyrGlyGluAspAspLeuAlaThrArg 292
DB 802 AACGGCTTCCCAACGATCGTGGGCTGGGCTGGCGAGATGACGATGCTGACAGG 861
QY 293 ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet 312
DB 862 GTCGCGCTGGCTGGAGTGAAGATCTCTGACACACTCACTCTGTGGGACACTTAAAGATG 921
QY 313 IleLysHisSerThrGlnAlaThrAsnProValAsnLysCysArgTyrLysIleMetGly 332
DB 922 GTGAGACACAGAGGGGATTAAGAAATGAGAAATCCACAGATTTGACCTTCCTGTC 981
QY 333 GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuTyrTyrLysLeuAlaAsn 352
DB 982 CGTACCCAGAAATCTTGACACAAAGATGAAATGAACTACCTACGATCCGACCTGCGCA 1041
QY 353 LeuGlnLysProLeuTyrThrArgAlaValAlaAspLeu 366
DB 1042 AGAAGCTGGGTCTCTCTATACCAACATCACTGCAGACATC 1083

RESULT 12
AL558425 970 bp mRNA linear EST 02-APR-2004
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL558425
VERSION AL558425.3 GI:46183823
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 970)
Li,W.B., Gruber,C., Jeesee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31280224.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
division of Invitrogen. This sequence belongs to sequence cluster
345.r
For more information about this cluster, see

```

```

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ005YK01"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 6,84e-71 Length: 970
Score: 680.50 Matches: 139
Percent Similarity: 62.00% Conservative: 47
Best Local Similarity: 46.33% Mismatches: 110
Query Match: 33.51% Indels: 5
DB: 1 Gaps: 3

US-10-661-430-1 (1-383) x AL558425 (1-970)
QY 64 AspAspLeuAspPheThrTrpAsnSerThrPheSerProIleSerGluVal---Aengln 82
DB 44 GATTCCTGTGGAGTGCCTGATGAGAGCTTTCTGCTGCTGCTGACATTTCCATT 103
QY 83 ThrSerPheMetGluAspIleArgProIleLeuPhe-ProAspAsnGlnThrLeuGlnPh 102
DB 104 AGATTAATGCTCAACATCTGTCGCCACAGGGGCTCTCCAGCTCTCAAGATGTCCTTA 163
QY 102 eCyAsnGlnThrProProHisLeuValGlyProIleArgVal---PheLeuAspGluTr 121
DB 164 CTGTCCAGAAACGATCTCTCTTACTGTGGCTGTGTGGTGTCTTTAGCCCAAGTCGC 223
QY 121 oAspPheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyLysIleGlyMetPr 141
DB 224 ATCATCGGCAGAGATTTGGAGCGGAATCCCGGATGAAACAGGGGCGCGATCCGCC 283
QY 141 oLysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAl 161
DB 284 TGCAGGTGTGAGCCCGCTCCGGAACGACATCATTTGCTCATGCTGACCGGAGCA 343
QY 161 aHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLeuGlnGlnLeuAspTyr 181
DB 344 CCACTGCGCTGCTCTCTTACCACTGCACCCCTTCTTGACGCGCCAGAGCTTCTTA 403
QY 181 fAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAs 201
DB 404 TGCATCTATGTATCATCCACAGGCTGGAATGAAATTTTHACAGGCAAAATCTTTGAA 463
QY 201 nValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVa 221
DB 464 CGTTGGGGTGCAGAGGCCCTGCGTGATGAAAGTGGAGCTGCTGTTCTTGACAGATGT 523
QY 221 lAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHis 240
DB 524 GGAACCTTGTCCAGAAATATGACCAATCTGTATGTGTGAGACCCCGGAGACCCGCCA 583
QY 240 sMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyG 260
DB 584 TGTTCGCTGTGCTATGAACAAGTTTGATACAGCTTCCGATCCCAAGTACTTCGGAGG 643
QY 260 yLysSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspPheTyr 280
DB 644 AGCTGACGACTTACTCTCCACAGTACCTGAAAGATGAATGCTTCCCAATGATATCTG 703
QY 280 pGlyTyrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysVa 300

```


Db 704 GGGCTGGGTGGTGAAGATGACACATTCTACACAGGCTGCGCTGGCTGGAGTGAAGAT 763
Qy 300 lserAgtTyrProthrGlnlleAlaArgTyrMetIlelybHisSerThrgluAlaThr 320
Db 764 CTCTGGCCCCCAGCATCTGTAGAGACATTAATAGAGACACCGAGCATGAAGAG 823
Qy 320 rAnpProValaenlybCyArGTrlyrlylleMetGlyInthrlybArGTrpThrAr 340
Db 824 CATGAGCAAAATCCCAAGACATTGA-CTTCTKAGCCCTTACCCAGAAATCTCGACGCA 882
Qy 340 gAapGlyLeuSerAenlybSerTyrlyrlybValaenlybGluLeuLybProleuTyr 359
Db 883 AGATGGAGTAAGTCACTGACATACAGTTGGTGGTGGAGCTGGGGCTCTTAT 940
RESULT 13
CX780748 791 bp mRNA linear EST 28-JAN-2005
LOCUS AGENCOURT 42763460 NIH_XGC_tropIncl Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7795167 5', mRNA sequence.
ACCESSION CX780748
VERSION CX780748.1 GI:58297541
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 791)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Bruce Blumberg
cDNA Library Preparation: B. Blumberg
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM15945 row: d column: 13
High quality sequence start: 13
High quality sequence stop: 714.
Location/Qualifiers
1. .791
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7795167"
/issue_type="Intestine"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH_XGC_tropIncl"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming
(5'-GAGAGAGAGAGAGAGAGACTAGTCTGAGTTTCTTTTCTTTTCTTTT-3')
and StrataScript reverse transcriptase. After ligation of
EcoRI adapters (5'-AATTCGACGACGAG-3') followed by kinasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested pCS107. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Michelle Tabb and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."

Alignment Scores:
Pred. No.: 2,326-70 Length: 791
Score: 675.00 Matches: 117
Percent Similarity: 68.70% Conservative: 52
Best Local Similarity: 47.56% Mismatches: 75
Query Match: 33.23% Indels: 2
DB: 8 Gaps: 1

US-10-661-430-1 (1-383) x CX780748 (1-791)

Qy 123 PheLySThrLeuGluLybIeTyrProAptHr-----HisAlaGlyGlyNIscLyMet 140
Db 9 TTGGCAGCATGATTCATTCATTCGACAGGCTTCGCCGCAAGAGGGCGCTCAAG 68
Qy 141 ProLybAapCyValAlaArgHisArgValAlaIleIleValProTyrArgAparGlu 160
Db 69 CCACCCGACTGGAGCTCCATCAATAAACGGCGGTGATATCCCTACCGCGGAGAG 128
Qy 161 AlaHisLeuArgIleMetLeuHisAenlybHisSerLeuLeuAlaLybGlnLeuAap 180
Db 129 CAGCACCTTAATAATACCTGCTGATATATCTGCACCCGTTTCTGCAGCGGACGCACTGAAC 188
Qy 181 TyrAlaIlePheIleValGluGlnValAlaAengInthrPheAenArgGlyLybLeuTet 200
Db 189 TACGGAGATATACATCATCATCATCGGGGTATTTACTTTCAACCGGCTAAAGTTGCTG 248
Qy 201 AenValIGlyTyrAparValAlaSerArgLeuTyrProTTPGInCybPheIlePheHisAap 220
Db 249 AACGTGGCTTCAGAGAGCGCATGAAGACGAGATGGAGCTGGTGTTCATCACAGAT 308
Qy 221 ValAapLeuLeuProGluAapAparArgAenlybTyrCybProIleInProArgHis 240
Db 309 GTGACCTCATATCCGGAAGAGATGCAATATCTATCTACCTGGAGCAAAATTTCCAAAGCAC 368
Qy 241 MetSerValAlaIleAapLybPheAenTyrlybLeuProTyrSerAlaIlePheGly 260
Db 369 GCCTGATCGCCATGACAGAAAGTTTGATTAACAATTCCTTCAAAATCTACTTCGGGGGA 428
Qy 261 IleSerAlaLeuThrLybAapHisLeuLybIleAengLybPheSerAapAparHeTTP 280
Db 429 GTGTAGCCCTTTCCTCCAGAACATATCAAGAGAAACGGCTTCCCAACAATCACTGG 488
Qy 281 GlyTTPGlyGlyLybAapAparAapLeuAlaThrArgThrSerMetAlaGlyLeuLybVal 300
Db 489 GGATGGCGCGCGGAAGATGATATATAGCATCAAGGTCGACCTTACGCGCATGATCATC 548
Qy 301 SerArgTyrProthrGlnlleAlaArgTyrlybMetIlelybHisSerThrgluAlaThr 320
Db 549 TCTCGCCCTTCTATCCACACACGAGGATATAAATGATCAAGCAGCGCATGACAAAGGCG 608
Qy 321 AenProValaenlybCyArGTrlyrlylleMetGlyInthrlybArgArgTrpThrArg 340
Db 609 AATGAGCAAAATCCCAAAAGGTTTAACATGCTGACGAAGACCGCGGACGTCGGCGAG 668
Qy 341 AapGlyLeuSerAenlybSerTyrlyrlybValaenlybGluLeuLybProleuTyr 360
Db 669 GACGGAGTAATCCCTGACAGTACTGCTCMTCAAGAGATTGACAGCGCTTTACACC 728
Qy 361 ArgAlaValaValaPleu 366
Db 729 AACATCATGTAACATC 746
RESULT 14
CX744001 803 bp mRNA linear EST 24-JAN-2005
LOCUS JGI ANBT1739.fwd NIH_XGC_tropA66 Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7514094 5', mRNA sequence.
ACCESSION CX744001
VERSION CX744001.1 GI:58070284
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ORIGIN

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 803)
Richardson, P., Lucas, S., Rokhsar, D., Dettler, J. C., Ng, D. C.,
Brokstein, P., and Lindquist, E. A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other ESTs: JGI ANBT1739.rev
Contact: Lindquist, E. A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
Clone Distribution: I.M.A.G.E. Consortium/LNLN:
<http://image.llnl.gov>
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: ANBT 0017 row: e column: 4
High quality sequence stop: 738.

FEATURES

source
1. 803
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7514094"
/tissue_type="whole embryo"
/dev_stage="Gastrula (st. 10-13)"
/lab_host="TOP10 E. coli"
/clone_1lb="NIH XGC tropGa66"
/note="Vector: PCS22+; Site 1: PstI; Site 2: XhoI. The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming (V(T)20CTCGAGAGAGAGAG). After ligation of PstI
adapters and by XhoI and PstI digestion, the cDNA was size
selected by chromatography on Sepharose CL-4B columns and
fractions containing cDNAs larger than 500 bp were ligated
into PstI/XhoI-digested PCS22+. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Jisong Peng and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."

ORIGIN

Alignment Scores:

Pred. No.: 2,38e-70 Length: 803
Score: 675.00 Matches: 115
Percent Similarity: 69.92% Conservative: 50
Best Local Similarity: 48.73% Mismatches: 71
Query Match: 33.23% Indels: 0
DB: 8 Gaps: 0

US-10-661-430-1 (1-383) x CX744001 (1-803)

QY 131 ProAaPThrHsAlaGlyGlyHsGlyMePProLysAspCyValAlaArgHisArgVal 150
Db 12 CTTACAGTGTCCAAAGAGGGCGCTACAAAGCACCCGAGCTCGAGTCACTCATTAACG 71
QY 151 AlaIleIleValPProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnLeu 170
Db 72 GCCGTGATCATCCCTTCACCGGCGAGGAGAGACCTTAATATCCGTCTGATTATCTG 131
QY 171 HisSerLeuLeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAla 190
Db 132 CACCCGTTTCTGACGGCGAGCACTGAACACGAGATATCATCATTCATCAGCGGGGT 191
QY 191 AsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeu 210

Db 192 AATTTCACCTTCACCGCGCTAAGTTGCAACGTCGCGCTTCACAGGAGCCCATGAAAGAC 251
QY 211 TyrProTgInGlyCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsn 230
Db 252 GAGGATTGGAGCTGTTGTTTACACAGATGTCAGTCTTATTCGCCAAGACATGCCAAT 311
QY 231 LeuTyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyr 250
Db 312 ATCTATACCTGGAGCAAAATTTCCCAAGACAGCGCTGATGCCATGAGCAAGTTGGATAC 371
QY 251 LysLeuProTyrSerAlaIlePheGlyIleSerAlaLeuThrLysAspHisIleuLys 270
Db 372 AAATTCCTTACAAATCTTACTTCGGGGAGAGTGTACGCCCTTTCGCCAACAATACATG 431
QY 271 LysIleAsnGlyPheSerAsnAspPheTyrGlyTyrGlyGluLysAspAspLeuAla 290
Db 432 AGATGACGGCTCTCCCAACAACTACTGGAGTGGGCGGAGGAGATGATGATATAGGC 491
QY 291 ThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyr 310
Db 492 ATCAGGGTGGCACTTACGGGCAATGATCATCTCGCCCTTATCCAGACGGAGATAT 551
QY 311 LysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIle 330
Db 552 AAATGATCAGACGCGCATGACAGGCGCAATGACAAATCCCAAAAGGTTTACATG 611
QY 331 MetGlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLysTyrLysLeu 350
Db 612 CTACAGAGACCGGGGAGCGTGGCGGAGAGGAGGATGAATCTCGTGCATCTACTG 671
QY 351 ValAsnLeuGluLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
Db 672 CTCTCAGAGAGTTGACGCGCTTTACACCAATCATCATGTGAACATC 719

RESULT 15

CX757659 796 bp mRNA linear EST 24-JAN-2005
LOCUS
AGENCOUNT 40973844 NIH MGC 281 Homo sapiens cDNA clone
IMAGE:7782009 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CX757659.1 GI:58054315
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Meri Firpo
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>

plate: LLAM15939 row: p column: 07
High quality sequence start: 16
High quality sequence stop: 792.
Location/Qualifiers

FEATURES

source

1. 796
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7782009"

/tissue_type="Pluripotent cell line derived from
blastocyst inner cell mass"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_281"
/note="Organ: Blastocyst; Vector: pExpress-1; Site: 1;
Score; Site 2: Not; RNA obtained from pluripotent cell
line derived from blastocyst inner cell mass (cell line
HSP-6, NIH Registry designation UC06. Positive for OCT4
expression by rtPCR, positive for SSEA-3, SSEA-4,
Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
SSEA-1 by immunofluorescence Passage 62. CDNA was primed
using oligo-dT primer:
5'-pGACTAGTCTAGATCGCAGCGCCGCC(7)25-3' and cloned into
the pExpress-1. Size-selection >1.25
kb resulted in an average insert size of 2.0 kb. This
primary library is normalized (non-normalized primary
library is NIH_MGC_280) and was constructed by Express
Genomics (Frederick, MD). Note: this is a Mammalian Gene
Collection library."

ORIGIN

Alignment Scores:
Pred. No.: 8.22e-70 Length: 796
Score: 670.50 Matches: 123
Percent Similarity: 66.12% Conservative: 39
Best Local Similarity: 50.20% Mismatches: 82
Query Match: 33.01% Indels: 1
DB: 8 Gaps: 1

US-10-661-430-1 (1-383) x CX757659 (1-796)

QY 117 PheLeuAspGluProAspPheIleuGluValIleTyrProAspThrHisAlaGly 136
DB 50 TTtACCCAGTGCATCTCACTGCAGAGATTGTGAGCGGAATCCCGGTAGAACCGAG 109
QY 137 GlyHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIleValProTyr 156
DB 110 GGGCCGGTACCGCCCTGCAGAGCTTGTGAGCCCGCTCCGAAACAGCATCATTTGGCTCAT 169
QY 157 ArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLys 176
DB 170 CGTGGCCGGGAGACACCTGCGCCCTCTTACCCACCTGACCCCTTCTTGCAGCGC 229
QY 177 GluGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArg 196
DB 230 CAGCAGCTTCTTATGGCATCTATGTCATCCACAGGCTGGAATGGAACTTTAAACAG 289
QY 197 GlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPhe 216
DB 290 GCAGAACTGTGAAGTTGGGGTGGAGAGGCGCTGCGATGAAGAGTGGAGACTGCTTG 349
QY 217 IlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys--Pro 235
DB 350 TTCTTGACAGATGTGACCTCTTGCAGAAATGACCAATCTGTATGTGTGACCC 409
QY 236 IleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSer 255
DB 410 CGGGGACCCCGGCATGTGCTGCTATGAACAGTTGAAACAGCTCCCGTACCCC 469
QY 256 AlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPhe 275
DB 470 CAGTACTTGGAGAGAGTCTCAGACACTTACTCTGACACAGTACTGAAAGATGAATGGCTTC 529
QY 276 SerAsnAspPheTyrGlyTyrGlyGluAspAspAspLeuAlaThrArgThrSerMet 295
DB 530 CCCAATGAATACTGGGGCTGGGGTGTGTGAGATGACGACATGTCTACCGGGTGGCGCTG 589
QY 296 AlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHis 315
DB 590 GCTGGGATGAAGATCTCTGCGCCCGCCACATCTGTAGACACTATAGAGATGGTGAAGCAC 649
QY 316 SerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGlyInThrLys 335
DB 335

DB 650 CGAGAGATAAGGCAATGAGGAAATCCACAGATTTGACCTTCGTGCTGATCCAG 709
QY 336 ArgArgTPTThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGlnLeu 355
DB 710 AATTCTGGACGACAGATGGGATGAACTCACTGACATTCAGATTGCTGGCTGAGAGCTG 769
QY 356 LysProLeuTyrThr 360
DB 770 GGGCTCTTTATACC 784

Search completed: January 9, 2006, 04:21:39
Job time : 4950 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 9, 2006, 00:54:55 ; Search time 236 Seconds
(without alignments)
2884.772 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSLLVLCV.....VDLEKDCRRLRRPFCF 383

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-O=/cg2_1/USPTO.spool/US10661430/runat_06012006_123136_6274/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humana4.0.cdi
-LIST=45 -DLOCALIGN=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10661430 @CGN1_1_193 @runat_06012006_123136_6274 -NCPU=6 -ICPU=3
-NO_MMAPP -LANG=ODDERY -NEG_SCORES=0 -WAIT -DSPBLCK=100 -LONCLIG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cg2_6/ptodata/1/ina/1 COMB.seq: *
2: /cg2_6/ptodata/1/ina/5 COMB.seq: *
3: /cg2_6/ptodata/1/ina/6A COMB.seq: *
4: /cg2_6/ptodata/1/ina/6B COMB.seq: *
5: /cg2_6/ptodata/1/ina/H COMB.seq: *
6: /cg2_6/ptodata/1/ina/PTUS COMB.seq: *
7: /cg2_6/ptodata/1/ina/PP COMB.seq: *
8: /cg2_6/ptodata/1/ina/RE COMB.seq: *
9: /cg2_6/ptodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	717	35.3	1938	3	US-09-673-395A-140
2	717	35.3	1970	3	US-09-949-016-2875
3	704	34.7	1189	3	US-09-270-767-11472
4	668.5	32.9	2304	2	US-08-446-777-5
5	666.5	32.8	2304	2	US-08-446-777-5
6	666	32.8	1265	2	US-08-446-777-1
7	658	32.4	1116	3	US-09-118-464-33
8	658	32.4	1116	3	US-10-132-652A-33
9	658	32.4	1119	3	US-09-949-016-2195

10	658	32.4	1119	3	US-09-949-016-2196	Sequence 2196, Ap
11	658	32.4	2027	3	US-09-118-464-1	Sequence 1, Appl
12	658	32.4	2027	3	US-10-132-652A-1	Sequence 1, Appl
13	646	31.8	1023	3	US-09-118-464-34	Sequence 34, Appl
14	646	31.8	1023	3	US-10-132-652A-34	Sequence 34, Appl
15	642.5	31.6	1964	3	US-09-991-181-235	Sequence 235, Ap
16	642.5	31.6	1964	3	US-09-990-444-235	Sequence 235, Ap
17	642.5	31.6	1964	3	US-09-997-333-235	Sequence 235, Ap
18	642.5	31.6	1964	3	US-09-992-598-235	Sequence 235, Ap
19	642.5	31.6	2151	3	US-09-949-016-1767	Sequence 1767, Ap
20	642.5	31.6	2167	3	US-09-949-016-541	Sequence 541, Ap
21	642.5	31.6	2280	2	US-09-055-097-4	Sequence 4, Appl
22	642.5	31.6	2280	3	US-09-373-902-4	Sequence 4, Appl
23	583	28.7	2965	3	US-09-949-016-4375	Sequence 4375, Ap
24	583	28.7	4646	3	US-09-949-016-680	Sequence 680, Ap
25	577	28.4	3830	3	US-09-949-016-679	Sequence 679, Ap
26	576	28.4	3832	3	US-09-949-016-2359	Sequence 2359, Ap
27	510.5	25.1	494	3	US-09-270-767-27061	Sequence 27061, A
28	446	22.0	541	3	US-09-404-879A-108	Sequence 108, Ap
29	446	22.0	541	3	US-09-338-933-108	Sequence 108, Ap
30	446	22.0	541	3	US-09-215-681-108	Sequence 108, Ap
31	446	22.0	541	3	US-09-216-003A-108	Sequence 108, Ap
32	446	22.0	541	3	US-09-667-857-108	Sequence 108, Ap
33	446	22.0	541	3	US-10-198-053-108	Sequence 108, Ap
34	446	22.0	541	3	US-09-827-271-108	Sequence 108, Ap
35	327.5	16.1	13290	3	US-09-949-016-13937	Sequence 13937, A
36	327.5	16.1	13290	3	US-09-949-016-13938	Sequence 13938, A
37	314.5	15.5	500	3	US-09-270-767-1933	Sequence 1933, Ap
38	314.5	15.5	500	3	US-09-270-767-1933	Sequence 1933, Ap
39	313	15.4	10383	3	US-09-949-016-14617	Sequence 14617, A
40	272.5	13.4	1524	3	US-10-012-231A-16	Sequence 16, Appl
41	272.5	13.4	1524	3	US-10-012-231A-16	Sequence 16, Appl
42	272.5	13.4	1524	3	US-10-015-389A-16	Sequence 16, Appl
43	272.5	13.4	1524	3	US-10-006-768A-16	Sequence 16, Appl
44	272.5	13.4	1524	3	US-10-015-671A-16	Sequence 16, Appl
45	272.5	13.4	1524	3	US-10-015-393A-16	Sequence 16, Appl
					US-10-011-833A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-673-395A-140
; Sequence 140, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 140
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-673-395A-140

Alignment Scores:

Pred. No.: 1.68e-82
Score: 717.00
Percent Similarity: 64.60%
Best Local Similarity: 48.91%
Query Match: 35.30%
DB: 3
Gaps: 2
US-10-661-430-1 (1-383) x US-09-673-395A-140 (1-1938)

Length: 1938

Matches: 134

Conservative: 43

Mismatches: 95

Indels: 2

[illegible][illegible]

```
RESULT 3
US-09-270-767-11472
; Sequence 11472, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11472
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11472

Alignment Scores:
Pred. No.: 3 64e-81 Length: 1189
Score: 704.00 Matches: 144
Percent Similarity: 61.49% Conservative: 46
Best Local Similarity: 46.60% Mismatches: 91
Query Match: 34.66% Indels: 28
DB: 3 Gaps: 7

US-10-661-430-1 (1-383) x US-09-270-767-11472 (1-1189)
QY 71 AanserThPhe---SerProIleSerGluValAsnGlnThrSerPheMetGluAspIle 89
DB 261 AACTCTACCATTAAGAACTGATGTCACAGGCACTTACTTCCATTCACAGACTTA 340
QY 90 ArgProIleLeu-----PheProAspAsnGln----- 98
DB 341 ACGCGCTTCCTGCGGCAAGAAATTTTCCCGCCGCAAGAAATCCATCCGCC 400
QY 99 ThrLeuGlnPheGysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeu 118
DB 401 CTCCTTCCCACTGCACTGATCCCGATCCCGTATGATGAGACCCATC----- 448
QY 119 ArgGluProAspPheLeuThrLeuGluValIle-----TyrPro 131
DB 449 ---AGGCCCAAC---ACGACACTGGAGTCACGACGTTATTGAGCGGACCTTGGACCT 502
QY 132 AspThrHisAlaGlyGlyHisLeuMetProGlyAspGysValAlaArgHisArgValAla 151
DB 503 CTTTTCGCCCGCTGGTGGCGCTTCCGAGAACTGCAATGCCGACATCAGCGTGGCT 562
QY 152 IleIleValProGlyArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHis 171
DB 563 ATTTGTTGCGCCCTCGCATGATGAGCCCATCTTACTTTCTCGCGCAACATCCAC 622
QY 172 SerLeuLeuAlaGlyGlnLeuAspGlyArgAlaIlePheIleValGluGlnValAlaAsn 191
DB 623 CCATTTCGTATGAGAGGCGCATCCGCTATCCATTTTCTATGTAGAGCAAGCAAGCGG 682
QY 192 GlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyr 211
DB 683 AAGCCCTTCATATCGGCGCTCCATGATGAACTTTGTTATTTGAGGCCCTTAAAGCTGTC 742
QY 212 ProThrGlnGysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeu 231
DB 743 CAGTGGAGATTGTTTATATTCACGATGTCGACTTCTTGGACGACGCAATCTT 802
QY 232 TyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLys 251
DB 803 TACAACCTGTCCACGTGACGCGGACACATGTCACTATGACACGCTGAACCTTCAAG 862
QY 252 LeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLys 271
DB 863 TTGCCTTATCATCATATTTTGGAGGTGTTTCCGCAATGACCGCTGAGACACTTTCAGGCC 922
```

```
QY 272 IleAsnGlyPheSerAsnAspPheTrpGlyTyrGlyGluAspAspLeuAlaThr 291
DB 923 GTAAATGGCTTCTCAAACTCCTTTGGCTGGCGCGCGAGATGACATGATGCCAAC 982
QY 292 ArgThrSerMetAlaGlyLeuLeuValSerArgTyrProThrGlnIleAlaArgTyrLys 311
DB 983 AGGTTGAAGCAGCGCCACACTTATTCATCAAGTATCCGATCAATGACCGCGCTACAG 1042
QY 312 MetIleValHisSerThrGluAlaThrAsnProValAlaLeuLysCysArgTyrLysIleMet 331
DB 1043 ATGCTGAAGCATCAGAGAGAAAGGCCAATCTTAAG-----CGCTATGAAGAACTTA 1093
QY 332 GlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuVal 351
DB 1094 CAGAAATGCATGACGAAATATGAAACAGATGGAATCACTGATTAATATATTCTATAC 1153
QY 352 AsnLeuGluLeuLysProLeuTyrThr 360
DB 1154 AGCATCAACAATTTCCAACTTTCAC 1180

RESULT 4
US-08-446-777-7
; Sequence 7, Application US/08446777
; Patent No. 5641668
; GENERAL INFORMATION:
; APPLICANT: Berger, Eric G.
; APPLICANT: Watzel, Manfred
; APPLICANT: Ivanow, Svetoslav X.
; TITLE OF INVENTION: Proteins having glycosyltransferase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferrari, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/B/BE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2304 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; STRAIN: E. coli DH5alpha
; IMMEDIATE SOURCE:
; CLONE: YEPGSTB
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2301
```


NAME/KEY: CDS
LOCATION: 1..2301
OTHER INFORMATION: /product=
OTHER INFORMATION: "galactosyltransferase-sialyltransferase hybrid
US-08-446-777-5
protein"

Alignment Scores:
Pred. No.: 8 48e-76 Length: 2304
Score: 666.50 Matches: 140
Percent Similarity: 54.188 Conservative: 74
Best Local Similarity: 35.448 Mismatches: 132
Query Match: 32.824 Indels: 49
DB: 2 Gaps: 6

US-10-661-430-1 (1-383) x US-08-446-777-5 (1-2304)

QY 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTyr----- 29
DB 76 CTGTCGCGCGCTGCGGCTGACCTTGCGCTGACCCCTGTTACTACCTGCGCGC 135
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
DB 136 GACCTGAGCGCGCTGCGCCCACTG-----GTGCGAGTCTCCACACCGCTG 180
QY 47 AlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
DB 181 CAGGCGCGCTCCAAACAGTCGCGCGCATCGGCGAGTCTCTCGGGAGACTCCGACCGGA 240
QY 61 -----SerThrSer 63
DB 241 GGGGCGCGCGCGCGCTCTCTAGGCGCGCTCTCCACCGCGCGCGGCGGCGAGCTCC 300
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
DB 301 AGCCAGTCGTGATTCGTGCGCGCGCGCTGAGCACTTACCTCGGCTC----- 351
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
DB 352 -----CGAGTCCCGCACACCGCACCTGCTGCCCGCTGC 390
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
DB 391 CCGTGAAGTCCCGCTGCTGCGCGCGCGCATGATGATTGAATTAACATGCTGTGAGC 450
QY 123 PheIysThrLeuGlnLysIleTyrProAspThrHisAlaGlyIleHisGlyMetProIys 142
DB 451 CTGAGAGCTGTGCAAGCAGAACCCCAATGTGAAGATGGCGCGCTATGCCCCAGG 510
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
DB 511 GACTGTGTCTCTCAAGGTGGCATCATCATTCCTCCAGAACCGGAGAGGAC 570
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrAla 182
DB 571 CTCAGTACTGCTATATATTATTGACCCAGTCTGTCAGCGCAGAGAGCTGAGACTATGCG 630
QY 183 IlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnVal 202
DB 631 ACTATGTATCAACACGCGGAGACACTATATTCATACGCTCAAGCTCTCATATGTT 690
QY 203 GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAsp 222
DB 691 GCGTTTCAAGAACCTTGAAGGACTATGACTACACTGCTTTGTGTGTTAGTACGTGGAC 750
QY 223 LeuLeuProGluAspAspArgAsnLeuTyrTrpCysProIleGlnProArgHisMetSer 242
DB 751 CTCATTCGAATGAATGACCATATATGCGTACAGTGTGTTTTCACAGCCAGCAGCATTTCC 810
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyGlyIleSer 262
DB 811 GTTGCAATGATTAAGTGTATTCAGCTTATGTTCAAGTATTTGAGAGTGTCTCT 870

QY 263 AlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTrpGlyTyr 282
DB 871 GCTCTAAGTAAACAACACTTTCTACCATCAATGATTTCCAAATAATATTGGGCGCTCG 930
QY 283 GlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
DB 931 GGAGAGAAAGATGATGATTTTAAAGATTAGTTTAAAGGAGCATGCTATATCTCGC 990
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
DB 991 CCAATTCGTGTGTGGAGGCTGTCATGATCCGCTCAAGACAGCAAGAAATGAA 1050
QY 323 ValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTyrThrArgAspGly 342
DB 1051 CCCATCTCAGAGGTTTGACCGAATTCACACACAAAGAGACAAATGCTCTGATGCT 1110
QY 343 LeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAla 362
DB 1111 TTGAATCACTCACTTACAGAGTGTGATGTAAGAGATACCCATTTGATATCCCAATTC 1170
QY 363 ValValAspLeuLeuGlnLysAspCysArgArgGluLeuArgArg 377
DB 1171 ACAGTGACATC-----GGACAGAGCTGGATCCGTGA 1206

RESULT 6
US-08-446-777-1
Sequence 1, Application US/08446777
Patent No. 5641668
GENERAL INFORMATION:
APPLICANT: Berger, Eric G.
APPLICANT: Matzele, Manfred
APPLICANT: Iwanow, Svetoslav X.
TITLE OF INVENTION: Proteins having glycosyltransferase
TITLE OF INVENTION: activity
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: CIBA-GEIGY Corporation
STREET: 520 White Plains Road
CITY: Tarrytown
STATE: NY
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.777
FILING DATE: May 26, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/03194
FILING DATE: 15 NOV 93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EPO 92810924.8
FILING DATE: 27 NOV 92
ATTORNEY/AGENT INFORMATION:
NAME: Ferrari, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 4-19361/A/BE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-3318
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
STRAIN: E. coli DH5alpha
IMMEDIATE SOURCE:


```

Db 256 -----CTGCCGGTCCCAAGGCTGCCAGCCCTGCTCTGA 296
QY nThrProH1sLeuVal1G1yPro1leArgVal1PheLeuAspGluPro---AspPhe1y 124
Db 297 CTGGCACTGCTGTGTGGGAGACTGCTGATCGAGTTCACTCAACCCATGCCCCCTGGA 356
QY 124 sThrLeuGlu1y1e1yPProAspThrH1sAlaG1y1H1sG1yMetPro1yAspCy 144
Db 357 GGGGGTGGAGAGGAGAACCCAGGCGTCTCATGGCGCGCATACACACGCCCGAGCTG 416
QY 144 sVal1AlaArgH1sArgVal1Ala1leVal1Pro1yArgAspArg1yAlaH1sLeu1x 164
Db 417 CACCCAGCCAGACGCTGGCTCATCTCCCTTTAACAACCCGGGAACACACCTGCG 476
QY 164 g1leMetLeuH1sAsnLeuH1sSerLeuLeuAla1y1eG1n1leuAsp1yAla1lePh 184
Db 477 CTACTGGCTCCCTCATCTTACACCCCATCTTGAAGCGGAGCGGCTGCGCTAGCGGCTCA 536
QY 184 e1leVal1G1n1Val1AlaAsnG1nThrPheAsnArg1y1y1eLeuMetAsnVal1G1y 204
Db 537 TGTCAATCAACAGCATGTGTAGAGAGACCTTCAACCGGCGCAAGCTGCTTAAGTGGGCTT 596
QY 204 r-----AspVal1AlaSerArgLeu1yPPro1yTrpG1nCyPhe1lePh 218
Db 597 CCTAGAGCGCTGAAGAGAGATGCCGCC-----TATGACTGCTCATCTT 641
QY 218 eH1sAspVal1AspLeuLeuProG1uAspAspArgAsnLeu1yTrH1yCyPPro1leG1Pr 238
Db 642 CAGCAGATGGAGACTGTGCTCCATGATGATACCGCAACCTATACCGCTGGCGGACCAACC 701
QY 238 oArgH1sMetSerVal1Ala1leAsp1yPheAsn1y1y1eLeuPro1ySerAla1lePh 258
Db 702 CGGCACATTTGCTGCTGATGCAATGGAAGATTGGCTTCCGCTTCCATGCTGCGCTACTT 761
QY 258 eG1yG1y1leSerAla1leuThr1yAspH1sLeu1yAla1y1eAng1yPheSerAsn1x 278
Db 762 TGGAGGTGTGTGAGGCTGTAGTAGGCTTCAGTAATCAATGAGCTTCCCAATGA 821
QY 278 pPheTrpG1yTrpG1yG1y1y1eAspAspAspLeuAlaThrArgThrMetAlaG1y1e 298
Db 822 GATACGTGGGCTGGGGTGGAGAGATGATACATCTTCAACCGGATCTCCCTACTGCGGAT 881
QY 298 u1ySerArg1yPProThrG1n1leAlaArg1y1y1eMet1y1eH1sSerThrG1 318
Db 882 GAAGATCTACGCCAGAGATCGAATCGCGCTACCGCATGATCAAGACAGACCGGCA 941
QY 318 uAlaThrAsnProVal1Asn1y1eCyArg1y1y1eMetG1y1nTh1yArgArgTr 338
Db 942 CAGCATTAACGAACCTTAACCTCAAGAGTTTACCAAGATTAACAAACACGAACTGACCAT 1001
QY 338 pThrArgAspG1y1eSerAsnLeu1yTr1y1y1eVal1AsnLeuG1uLeu1yPPro1e 358
Db 1002 GAAGCGGAGCGGCTGCTGAGTCAAGTCAAGGCTTGAAGGTGTCTCGGCAACCAT 1061
QY 358 uTyTrH1sArgAlaVal1AspLeu 366
Db 1062 CTTCACCAATATACAGATGACATT 1086

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652A-33

Alignment Scores:
Pred. No.: 3,2e-75 Length: 1116
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
DB: 3 Gaps: 6

US-10-661-430-1 (1-383) x US-10-132-652A-33 (1-1116)
QY 7 AlAVal1AlaArgLeu1y1ySerLeuLeuVal1leuCyAlaVal1leuLeuVal1H1sAl 26
Db 24 GCTGAGCGCGCTGTGCAAGCGCTGTCTCTTCTGCTGCTGCACTTCTGTGCGCT 83
QY 26 aMet1y1y1y1ePProSerLeu1yG1uAsnLeuThr1leG1ySerSer---Th1e 45
Db 84 CATCTCTAC---TTGAGCTTACGCCAGACACCTGAGCTTCTTCAAGCCCTTCAGTGC 140
QY 45 u1leAlaAspVal1AspAlaMetG1uAlaVal1leuG1yAsnThrAlaSerThrSerAspAs 65
Db 141 CCGAGGCCCTGCCATGCCCTTCCACCCAGCTGTGACAGCAGCAGACAGACGACGACTG 200
QY 65 p1eLeuAspThrThrAsnSerThrPheSerPro1leSerG1uVal1AsnG1nThrSerPh 85
Db 201 CTCCCGGCCCAACGCCACCGCTGTAGCTCCGGGCTCCCTGAGGTCCCAAGTCC----- 255
QY 85 eMetG1uAsp1leArgPro1leLeuPheProAspAsnG1nThrLeuG1nPh1eCyAsnG1 105
Db 256 -----CTGCCGGTCCCAAGGCTGCCAGCTGCCAGCTGCCACCTGTCTGA 296
QY 105 nThrProProH1sLeuVal1G1yPro1leArgVal1PheLeuAspGluPro---AspPhe1y 124
Db 297 CTGGCACTGCTGTGTGGGAGACTGCTGATCGAGTTCACTCAACCCATGCCCCCTGGA 356
QY 124 sThrLeuGlu1y1e1yPProAspThrH1sAlaG1y1H1sG1yMetPro1yAspCy 144
Db 357 GGGGGTGGAGAGGAGAACCCAGGCGTCTCATGGCGCGCATACACACGCCCGAGCTG 416
QY 144 sVal1AlaArgH1sArgVal1Ala1le1leVal1Pro1yArgAspArg1yAlaH1sLeu1x 164
Db 417 CACCCAGCCAGACGCTGGCGGTCTCATCTCCCTTTAGACACCGGGAACACACCTGCG 476
QY 164 g1leMetLeuH1sAsnLeuH1sSerLeuLeuAla1y1eG1n1leuAsp1yAla1lePh 184
Db 477 CTACTGGCTCCCTCATCTTACACCCCATCTTGAAGCGGAGCGGCTGCGCTAGCGGCTCA 536
QY 184 e1leVal1G1n1Val1AlaAsnG1nThrPheAsnArg1y1y1eLeuMetAsnVal1G1y 204
Db 537 TGTCAATCAACAGCATGTGTAGAGAGACCTTCAACCGGCGCAAGCTGCTTAAGTGGGCTT 596
QY 204 r-----AspVal1AlaSerArgLeu1yPPro1yTrpG1nCyPhe1lePh 218
Db 597 CCTAGAGCGCTGAAGAGAGATGCCGCC-----TATGACTGCTCATCTT 641
QY 218 eH1sAspVal1AspLeuLeuProG1uAspAspArgAsnLeu1yTrH1yCyPPro1leG1Pr 238
Db 642 CAGCAGATGGAGACTGTGCTCCATGATGATACCGCAACCTATACCGCTGGCGGACCAACC 701
QY 238 oArgH1sMetSerVal1Ala1leAsp1yPheAsn1y1y1eLeuPro1ySerAla1lePh 258
Db 702 CGGCACATTTGCTGCTGATGCAATGGAAGATTGGCTTCCGCTTCCATGCTGCGCTACTT 761
QY 258 eG1yG1y1leSerAla1leuThr1yAspH1sLeu1yAla1y1eAng1yPheSerAsn1x 278
Db 762 TGGAGGTGTGTGAGGCTGTAGTAGGCTTCAGTAATCAATGAGCTTCCCAATGA 821

```

Qy	278	p h e t r i p g l y t r p g l y g l u a s p a s p a s p l e u a l a t h r a r g h t s e r m e t l a g l y l e	298
Db	822	g t a c t g g g g c t g g g g t g c g a g a g a t a c t t t c a a c c g a t c t c c t g a c t g g a t	881
Qy	298	u l y e v a l s e r a r g t y r p r o t h i n g l i l a l a a r g t y l y s m e t i l e l y s h i s e r t h g l	318
Db	882	g a a g a t c t c a c g c c c a g a c a t c c c a a t c g c c g c t a c c g a t g a t c a a g c a c g a c c g c g a	941
Qy	318	u a l a t h r a s n p r o v a l a e n l y s c y s a r g t y l y s l l e m e t g l y n t h r l y s a l g a r g t r	338
Db	942	c a a g c a t a t a c g a a c c t a a c c c t c a g a g g t t t t a c a a g a t t c a a a c a c a g a a g c t g a c c a t	1001
Qy	338	p t h r a r g a s p g l y l e u s e r a s m l e u l y s t y r l y s l e u a l a s n d e u g l u l e u y s p r o l e	358
Db	1002	g a a g c g g a c c g c a t t g g g t c a g t c g c g g a c a c a g s t c t t t g a a g g t c t c g c g a a c c a c t	1061
Qy	358	u t y r t h r a l a l a v a l a s p l e u	366
Db	1062	c t t c a c a a t a t c a c a g t g a c a t t	1086

RESULT 9
US-09-949-016-2195
; Sequence 2195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

US-10-661-430-1 (1-383) x US-09-949-016-2195 (1-1119)
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 2195
/ LENGTH: 1119
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-2195

Alignment Scores:
Pred. No.:      3.21e-75      length:      1119
Score:          658.00        Matches:     139
Percent Similarity: 56.91%    Conservative: 71
Best Local Similarity: 37.67%  Mismatches:  136
Query Match:    32.40%      Indels:      23
DB:             3           Gaps:        6

```

```

QY      7 AAlaValAlaArgLeu-LysSerLeuLeuValLeuCysAlaValLeuLeuValHisAl 26
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      24 GGTGAGAGCCGCTGTGCAGAGCTGTGCTCTTCTGCTGCTGACACTCTTCGTGGCCGT 83
QY      26 aMetIleTyTyrLysIleProSerLeuTyrgluAnleuThrIleGlySerSer---ThrIle 45
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      84 CATTCCTTCAC--TTTGAGCTCTAGGCCACGACCTGGCTCTTACGCGCTTCACGTGC 140
QY      45 uIleAlaAspValAspAlaMetGluAlaValLeuGluGlyAsnThrAlaSerThrSerAspAs 65
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      141 CCGAGGCCCGCTGCACATGCTCCCTCACCCACGCTGTAGACGACGACGACGACGACACTG 200
QY      65 pIueuLeuAspThrTyrPasnSerThrPheSerProIleSerGluValAsnGlnThrSerPh 85
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      201 CTCGCCGACCCACGCCACGCGCTCTGACTCGGAGCTCTCGAGAGTCCCAAGTGC----- 255
QY      85 eMetGluAspIleArgProIleLeuPheProPheAsnGlnThrLeuGlnPheCysAsnGln 105
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

[illegible]

```

RESULT 10
US-09-949-016-2196
/ Sequence 2196, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/

```



```
QY 105 nThrProtonHleuValGlyProIleArgValPheIleuAspGluPro---AspPheIu 124
DB 466 CTGCGCACTGGCTGTGTGGGAGAGACTGGATGAGTTGACCTGACCTGACCGGCGGCGA 525
QY 124 vThrLeuGluValIleTyProAspThrThiValAGlyGlyHsiGlyMetProIuAspCy 144
DB 526 GCGGGTGCAGAGAGAGAACCCAGCGGTGCTCATGAGGCGCGCGATACACACCGCGCGACTG 585
QY 144 vValAlaArgHsiArgValAlaIleIleValProTyArgAspArgGluAlaHsiIleuTy 164
DB 586 CACCCCAAGCCCAAGCGGTGAGTATCCCTTTAGACCGGAGACACACACCACTGGCG 645
QY 164 gIleMetLeuHsiAsnLeuHsiSerLeuLeuAlaIuysGlnGlnLeuAspTyAlaIlePh 184
DB 646 CTACTGGCTTCATATCTTACACCCCACTTGAGCGCGCGACGCGCTGGCGTACGCGCTTA 705
QY 184 eIleValAGluGlnValAlaAsnGlnThrPheAsnArgGlyIuysLeuMetAsnValGlyTy 204
DB 706 TGTCTATCACACAGCATGTGAGGACACTTCAACCGGCGCACAGCTGCTTAACTGGGCTT 765
QY 204 r-----AspValAlaSerArgLeuTyProIleGlnCyAspHeIlePh 218
DB 766 CTTAGAGCGCGCTAAGAGAGATGCCGCC-----TATGACTGCTTCACTT 810
QY 218 eHsiAspValAspLeuProGluAspAspArgAsnLeuTyThrCyProIleGlnPr 238
DB 811 CACGAGATGTGACTGGTCTCCCATGATGATGACCGGAACCTTATCCGCGCGGCGACCAACC 870
QY 238 oArgHsiMetSerValAlaIleAspIuysPheAsnTyIuysLeuProTySerAlaIlePh 258
DB 871 CCCCACCTTGTCCATGTCATGACAGTGTGGCTTCCGCGCTTCCCTATCTGGCTACTT 930
QY 258 eGlyGlyIleSerAlaIleuThiIuysAspHsiIleuValIleAsnGlyPheSerAsnAs 278
DB 931 TGGAGGTGTGTACAGCGCTGAGTACAGCTCTTCTAGATATCATGCTTCCCACTGA 990
QY 278 pHeTrpGlyTyIleGlyIuysAspAspLeuAlaIleThiArgThrSerMetAlaGlyLe 298
DB 991 GTATCGGGGTGGGTGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
QY 298 uIuysValSerArgTyProThrGlnIleAlaGlyTyIuysMetIleIuysHsiSerThrG1 318
DB 1051 GAAGATCTCACGCCACAGACATCCCAATGCGCGCTTACCGCATGATCAAGACGACCGCGA 1110
QY 318 uAlaIleThiAspProValAsnIuysCyAspGlyIuysIleMetGlyGlnThiIuysArgTr 338
DB 1111 CAAACATTAAGCAACCTTAACCTCAGAGGTTTACCAAGATTCAAAACAGAAAGCTGACCAT 1170
QY 338 pThiArgAspGlyLeuSerAsnLeuTyIuysIleValAsnLeuGluIleuTyAspProLe 358
DB 1171 GAAGCGGAGACGCAATTGGGTACAGTGGCGTACAGGTCTTGGAGGTGTCTGGGCAACCACT 1230
QY 358 uTyTrpThiArgAlaValAlaIleu 366
DB 1231 CTTACACCAATATCACAGTGGACATT 1255
RESULT 12
US-10-132-652A-1
; Sequence 1, Application US/10132652A
; Patent No. 6916649
; GENERAL INFORMATION:
; APPLICANT: CLAUSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; FILE OF INVENTION: 4-GALACTOSYLTRANSFERASE, B4 GAL-72
; CURRENT APPLICATION NUMBER: US/10/132, 652A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
```

```
; LENGTH: 2027
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652A-1
Alignment Scores:
Pred. No.: 8,74e-75
Score: 658.00
Percent Similarity: 56.91%
Best Local Similarity: 37.67%
Query Match: 32,408
DB: 3
Gaps: 6
US-10-661-430-1 (1-383) x US-10-132-652A-1 (1-2027)
QY 7 AlaValAlaArgLeuIleuSerLeuLeuValIuysCyAsnValAlaIleuLeuValAlaIle 26
DB 193 GCTGAGCGCGCTGTGAGAGCTGTGCTCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
QY 26 aMetC1eTyIuysIleProSerLeuTyGluAsnLeuThiIleGlySerSer---ThiLe 45
DB 253 CATCTCTAC---TTGAGCTTACGCCACGACGACCTGGCGCTTCTTACGCGCTTACGCTG 309
QY 45 uIleAlaAspValAspAlaMetC1uAlaValIleuGlyAsnThiAlaSerThrSerAspAs 65
DB 310 CCGAGGCGCTGCCATGCGCTTCCACCCAGCTGACACAGCAGCAGCAGCAGCAGCAGCAGCAG 369
QY 65 pIleuAspThiTrpAsnSerThrPheSerProIleSerGlyValAlaAsnGlnThiSerPh 85
DB 370 CTCCCGGCCCAACGCCCGCTTACGCTCCGAGCTCCCTGAGTCCCAAGTCCCAGTCCC 424
QY 85 eMetGluAspIleArgProIleuPheProAspAsnGlnThiLeuGlnPheCyAsnG1 105
DB 425 -----CTGCCGCTCCCAACGCCCGCTCCCAAGCTCCCAAGCTCCCAAGCTCCGA 465
QY 105 nThrProtonHsiLeuValGlyProIleArgValPheIleuAspGluPro---AspPheIu 124
DB 466 CTGCGCACTGGCTGTGTGGGAGAGACTGGATGAGTTGACCTGACCTGACCGGCGGCGA 525
QY 124 vThrLeuGluValIleTyProAspThrThiValAGlyGlyHsiGlyMetProIuAspCy 144
DB 526 GCGGGTGCAGAGAGAGAACCCAGCGGTGCTCATGAGGCGCGCGATACACACCGCGCGACTG 585
QY 144 vValAlaArgHsiArgValAlaIleIleValProTyArgAspArgGluAlaHsiIleuTy 164
DB 586 CACCCCAAGCCCAAGCGGTGAGTATCCCTTTAGACCGGAGACACACCACTGGCG 645
QY 164 gIleMetLeuHsiAsnLeuHsiSerLeuLeuAlaIuysGlnGlnLeuAspTyAlaIlePh 184
DB 646 CTACTGGCTTCATATCTTACACCCCACTTGAGCGCGCGACGCGCTGCGCTTACGCGCTTA 705
QY 184 eIleValAGluGlnValAlaAsnGlnThrPheAsnArgGlyIuysLeuMetAsnValGlyTy 204
DB 706 TGTCTATCACACAGCATGTGAGGACACTTCAACCGGCGCACAGCTGCTTAACTGGGCTT 765
QY 204 r-----AspValAlaSerArgLeuTyProIleGlnCyAspHeIlePh 218
DB 766 CTTAGAGCGCGCTAAGAGAGATGCCGCC-----TATGACTGCTTCACTT 810
QY 218 eHsiAspValAspLeuProGluAspAspArgAsnLeuTyThrCyProIleGlnPr 238
DB 811 CACGAGATGTGACTGGTCTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
QY 238 oArgHsiMetSerValAlaIleAspIuysPheAsnTyIuysLeuProTySerAlaIlePh 258
DB 871 CCCCACCTTGTCCATGTCATGACAGTGTGGCTTCCGCGCTTCCCTATCTGGCTACTT 930
QY 258 eGlyGlyIleSerAlaIleuThiIuysAspHsiIleuValIleAsnGlyPheSerAsnAs 278
DB 931 TGGAGGTGTGTACAGCGCTGAGTACAGCTTCTTGAAGATCAATGAGGTCTCCCACTGA 990
QY 278 pHeTrpGlyTyIleGlyIuysAspAspLeuAlaIleThiArgThrSerMetAlaGlyLe 298
DB 991 GTATCGGGGTGGGTGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
```



```
Db 991 GTAGCTGGGCTGGGGTGGCGAGATGATGACATCTTCAACCGGATCTCCCTGACTGGGAT 1050
Qy 298 ulysValSerArgTyrProThrGlnlleAlaArgTyrIysMetIleIysHisSerThrl 318
Db 1051 GAAGATCTCAACCGCCAGACATCCGATCGCGCTACCGCATGATCAAGACAGACCGCGA 1110
Qy 318 uAlaThrAsnProValAsnIlyeCyAsrGlyIysIleMetGlyGlnThrIysArgArgTr 338
Db 1111 CAAAGATACGAAACCTAACCCCTCAGAGCTTTCAAAGATTCAAACAGACAGACCTGACAT 1170
Qy 338 pThrArgAspGlyIysSerAsnIleuIystrIyIysIleuValAsnIleuGluIleuIysProle 358
Db 1171 GAAGCGGAGCGGCTGAGTGGGTGAGTGGGTGAGTGGGTGAGTGGGTGAGTGGGTGAGT 1230
Qy 358 uTyrThrArgAlaValAlaAspIleu 366
Db 1231 CTTCAACCAATATCAAGTGGACATT 1255

RESULT 13
US-09-118-464-34
; Sequence 34, Application US/09118464A
; Patent No. 6558934
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
; TITLE OF INVENTION: Beta-1,4-Galactosyltransferase, Beta4Gal-T2
; FILE REFERENCE: 4305/08521
; CURRENT APPLICATION NUMBER: US/09/118,464A
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-118-464-34

Alignment Scores:
Pred. No.: 1,01e-73 Length: 1023
Score: 646.00 Matches: 120
Percent Similarity: 61.65% Conservative: 52
Best Local Similarity: 43.01% Mismatches: 95
Query Match: 31.81% Indels: 12
DB: 3 Gaps: 3

US-10-661-430-1 (1-383) x US-09-118-464-34 (1-1023)
Qy 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisIleuValAlaIysProIle 114
Db 172 CCCAGGCTCCCAACGCTGACCCCTGCTGACTGCGCCACCTGCTTGTGGGCAAGACTG 231
Qy 115 ArgValIleLeuAspGluPro--AspPheIysThrLeuGluIlyIleTyrProAspThr 133
Db 232 CTGATCGAGTTCACCTCAACCTCAACGCGCGTGCAGAGGGAAGAACCCAGGCGTG 291
Qy 134 HisAlaGlyGlyHisGlyMetProIyAspGlyValAlaArgHisArgValAlaIleIle 153
Db 292 CTCATGGGCGCGGATACACACCGCCGCTGACCTGACCCGACGAGCGTGGCGGTCATC 351
Qy 154 ValProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnIleuHisSerIleu 173
Db 352 ATCCCTTTAGACACCGGAGACACACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411
Qy 174 IeuAlaIysGlnGlnIleuAspTyrAlaIlePheIleValGlyGlnValAlaAsnGlnThr 193
Db 412 TTGAGGCGGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
Qy 194 PheAsnArgGlyIysIleuMetAsnValGlyTyr-----AspValAla 207
Db 472 TTCACACCGGCGGAGGCTGTTAACGTCGCTTCTTACGAGGCGCTGAAGAGGATGCGGCG 531
Qy 208 SerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAspIleuIleuProGluIleu 227
```

```
Db 532 -----TATGACTGCTTCATCTTCAAGCGATGAGACTGGTCCCATGAT 576
Qy 228 AspArgAsnIleuTyrThrCysProIleGlnProArgHisMetSerValAlaIleIysIys 247
Db 577 GACCGCAACCTATACCGCTGGCGGACACCGCCGCACTTGGCCATTGCAATGGAACAAG 636
Qy 248 PheAsnTyrIysLeuProTyrSerAlaIlePheGlyIlyIleSerAlaIleuThrIysAsp 267
Db 637 TTGGCTTCGCGCTTCCTATGCTGCGCTACTTGGAGGATGATGAGGCTGAGTAAAGCT 696
Qy 268 HisIleuIyIysIleAsnGlyPheSerAsnAspPheIleIyIyIyIyIyIyIyIyIyIyIy 287
Db 697 CAGTTTCAGATCAATAGGCTTCCCAATAGTACTGGGCTGGGCGTGGGAGAGATGAT 756
Qy 288 AspIleuAlaThrArgThrSerMetAlaGlyIleuIysValSerArgTyrProThrGlnIle 307
Db 757 GACATCTTCAACCGGATCTCCCTGACTGGATGAAGATCTCAACGCCAGACATCCGAATC 816
Qy 308 AlaArgTyrIysMetIleIysHisSerThrGlnAlaThrAsnProValAsnIysCysArg 327
Db 817 GCGCGCTACCGATGATCAAGACACCGCGCAAGCATTAACGAACCTAACCTTACAGAG 876
Qy 328 TyrIyIysIleMetGlyGlnThrIysArgArgTrpThrArgAspGlyIysSerAsnIleuIys 347
Db 877 TTTACCAAGATTCAAAACACGAGCTGACATGAAGCGGAGCGGATGGGTGAGTGGCGG 936
Qy 348 TyrIyIysIleuValAsnIleuGluIleuIysProIleuTyrThrArgAlaValAlaAspIleu 366
Db 937 TACCAAGTCTTGGAGTGTCTCGGCAACACTCTTCAACCAATATCAAGTGGACATT 993

RESULT 14
US-10-132-652A-34
; Sequence 34, Application US/10132652A
; Patent No. 6916649
; GENERAL INFORMATION:
; APPLICANT: CLAUSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; TITLE OF INVENTION: 4-GALACTOSYLTRANSFERASE, B4 GAL-T2
; FILE REFERENCE: 04305/100521-US1
; CURRENT APPLICATION NUMBER: US/10/132,652A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652A-34

Alignment Scores:
Pred. No.: 1,01e-73 Length: 1023
Score: 646.00 Matches: 120
Percent Similarity: 61.65% Conservative: 52
Best Local Similarity: 43.01% Mismatches: 95
Query Match: 31.81% Indels: 12
DB: 3 Gaps: 3

US-10-661-430-1 (1-383) x US-10-132-652A-34 (1-1023)
Qy 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisIleuValAlaIysProIle 114
Db 172 CCCAGGCTCCCAACGCTGACCCCTGCTGACTGCGCCACCTGCTTGTGGGCAAGACTG 231
Qy 115 ArgValIleLeuAspGluPro--AspPheIysThrLeuGluIlyIleTyrProAspThr 133
Db 232 CTGATCGAGTTCACCTCAACCTCAACGCGCGTGCAGAGGGAAGAACCCAGGCGTG 291
Qy 134 HisAlaGlyGlyHisGlyMetProIyAspGlyValAlaArgHisArgValAlaIleIle 153
Db 292 CTCATGGGCGGATACACACCGCCGCTGACCTGACCCGACGAGCGTGGCGGTCATC 351
Qy 154 ValProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnIleuHisSerIleu 173
Db 352 ATCCCTTTAGACACCGGAGACACACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411
Qy 174 IeuAlaIysGlnGlnIleuAspTyrAlaIlePheIleValGlyGlnValAlaAsnGlnThr 193
Db 412 TTGAGGCGGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
Qy 194 PheAsnArgGlyIysIleuMetAsnValGlyTyr-----AspValAla 207
Db 472 TTCACACCGGCGGAGGCTGTTAACGTCGCTTCTTACGAGGCGCTGAAGAGGATGCGGCG 531
Qy 208 SerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAspIleuIleuProGluIleu 227
```

```
Db 292 CTCATGGGCGCGCATACACACCCGCGATCGACCCGACGCGTGGCGGTCAAC 351
Qy 154 ValProDyRArgspAhrGua1aH1sleuArg1L1emleuH1saenueH1sser1eu 173
Db 352 ATCCCTTTAGACACCGGAAACACACTGCTGCTACTGGCTGCTACATCTACACCCCATC 411
Qy 174 LeuAlaLysGlnInleuApyrAla1lePhe1Leu1Glu1Gln1Val1AsnGlnThr 193
Db 412 TTGAGCGCGGAGCGGCTCGGCTACGCGGTCTATGTCATCAACGACGATGGTAGACACC 471
Qy 194 PheAsnArgGlyLysleuMetAsnValGlyTyR-----AspValAla 207
Db 472 TTCACCGCGGCGGCTGCTTAACTGGGCTTCTAGAGCGCGGTGAAGAGATGCCGCC 531
Qy 208 SerArgLeuTyRProTrpGlnCyRphe1LePheH1saPvalAspLeuPProGluAsp 227
Db 532 -----TATGACTGCTTCATCTTCAAGCATGAGACTGGTCCCATGAT 576
Qy 228 AspArgAsnLeuTyRThrCyRProL1egInProArgH1sMetSerValAla1leApyrLys 247
Db 577 GACCGCAACTATACCGCTGCGGACCAACCCGCACTTGGCATTTGCCATGGACAGAAC 636
Qy 248 PheAsnTyRlysLeuProTyRserAla1lePheGlyGlyLysSerAlaLeuThrLysAsp 267
Db 637 TTGGGCTTCCGCGCTCCCTATGCTGCTGCTACTTGGAGGTGTGTACAGGCTGAGTAAAGGCT 696
Qy 268 H1leuLysLysL1eAsnGlyPheSerAsnAspPheTrpGlyTyRTrpGlyGlyLysAsp 287
Db 697 CAGTTCTGGAATCAATGAGCTTCCCATAGAGACTGGGAGTGGGAGGAGAGAT 756
Qy 288 AspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyRProThrGln1le 307
Db 757 GACATCTTCAACCGGATCTCCCTGACATGGAGTGAAGATCTCACGCCACATCCGATC 816
Qy 308 AlaArgTyRlysMetL1elysH1serThrGluAlaThrAsnProValAsnLysCyRArg 327
Db 817 GGGCGCTTACCGCATGATCAAGACGACCGGACAGCATACGAACCTTAACTTCAGAGG 876
Qy 328 TyRlysL1eMetGlyInThrLysArgArgTyRThrArgAspGlyLysSerAsnLeuLys 347
Db 877 TTATACCAAGATTCAAAACAGAAAGCTTACCATGAAGGAGGAGGATTTGGGTCAGTCCG 936
Qy 348 TyRlysLeuValAsnLeuGluLeuLysProLeuTyRThrArgAlaValLysAspLys 366
Db 937 TACAGGCTTGGAGGTGTCTCGGCAACCACTTTCACCAATATCAACAGTGGACATT 993

RESULT 15
US-09-991-181-235
/ Sequence 235, Application US/09991181
/ Patent No. 6913919
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary B.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: K1javin, Ivar J.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: P2730PTC53
/ CURRENT APPLICATION NUMBER: US/09/991,181
/ CURRENT FILING DATE: 2001-11-16
/ PRIOR APPLICATION NUMBER: 60/049787
/ PRIOR FILING DATE: 1997-06-16
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/065186
/ PRIOR FILING DATE: 1997-11-12
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066770
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/075945
/ PRIOR FILING DATE: 1998-02-25
/ PRIOR APPLICATION NUMBER: 60/078910
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/084600
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/087106
/ PRIOR FILING DATE: 1998-05-28
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087609
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087759
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087827
/ PRIOR FILING DATE: 1998-06-03
/ PRIOR APPLICATION NUMBER: 60/088021
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088025
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088026
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088028
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088029
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088030
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088033
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088326
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088167
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088202
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088212
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088217
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088655
/ PRIOR FILING DATE: 1998-06-09
/ PRIOR APPLICATION NUMBER: 60/088734
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088738
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088742
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088810
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088824
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088826
/ PRIOR FILING DATE: 1998-06-10
```

PRIOR APPLICATION NUMBER:	60/088858
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/088651
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/088767
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/089105
PRIOR FILING DATE:	1998-06-12
PRIOR APPLICATION NUMBER:	60/089440
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089512
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089514
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089522
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089538
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089598
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089599
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089600
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089653
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089801
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089908
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089948
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089952
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090246
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090252
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090254
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090349
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090355
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090429
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090411
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090435
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090444
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090445
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090472
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090535
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090540
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090540
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090542
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090557
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090676
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090678
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090690
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090690
PRIOR FILING DATE:	1998-06-25

Query	Subject	Score	Length	Matches	Mismatches	Indels	Gaps
US-10-661-430-1 (1-383) x US-09-991-181-235 (1-1964)		8.65e-73	1964	143	55	124	8
QY	HisLeuAlaValAlaLAAArgLeuSerLeuLeuValLeuCyValAlaValLeuLeu	642.50	143	55	124	8	
DB	140 CACCTTCTCTAAACAAATTCGATTAATCTGCTGTGACTTGTCTCCGACAGTGGTGGG	53.37%	55	124	8		
QY	23 -----LeuValHisAlaMetIleTyrValIleProSerLeuTyrGlu	31.63%	49	8			
DB	200 TGGGCACACAGTAACACTGCTGGTGGTCC--ATTCAAGAGATCTCTAAAGCAAGAG	31.63%	49	8			
QY	37 AsnLeuThrIleGlySerSerThrLeuIleAlaSerValAlaSerAlaMetGluAlaValLeu						
DB	257 TTCATGGCTAAATTTCCATTAAGACCTCATTTTG-----GGAAAGCGAAAATCTGTG						
QY	57 GlyAsnThrAlaSerThrSerAspAspLeuAspThrTrpAsnSerThrPheSerPro						
DB	308 ACTAATGAAGATCCACCAAGAAAGTAGAAGTTGCAAC-----						
QY	77 IleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAsp						
DB	346 -----						
QY	97 AsnGlnThrLeuGlnPheCyAsnGlnThrProProHisLeuValGlyProIleArgVal						
DB	347 -----TGTCCTTCTGTGTCTTCACTTCAAGGCGCAGAGAACTC						
QY	117 PheLeuAspGlnProAspPheLeuThrLeuGluValIleTyrProAspThrHis						
DB	389 ATTTCG--AAACCAAGATCTC--ACTTGGAAAGAGTACAGCAGAAATCCCAAATG						
QY	135 AlaGlyGlnHisGlyMetProLeuAspSerValAlaArgHisArgValAlaIleIleVal						
DB	443 TCCAGAGGCGGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCCGCAATCTGTT						
QY	155 ProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisIleuLeuHisSerLeuLeu						

```
Db      |||::|||::||| ||| |||::||| ||| |||
503 CCCACCGGAAACAGAGAAACATGATGATCTGCGAATCTGCATCCCTTCCTG 562
QY      175 AlAlvGInGInleuAspTYrAlAlIephelIeValGInGInValAlaAsnGInThrPhe 194
Db      563 CAGAGGCAGCAGCTGATTTAGTCATCTACGATCCACCGAGCTGAAGGTAAAGATT 622
QY      195 AsnArgGlyLysLeuMetAsnValGlyTYrAspValAlaSerArgLeuTYrProTyrGln 214
Db      623 AATCGAGCCAAACTTGAATGTGGCTATCTAGAACCTCTCAAGAGAAATTTGGAC 682
QY      215 CysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTYrThrCys 234
Db      683 TGCTTATATATCCACGATGTGACCTGGTACCCGAGATGACTTTAACTTTACAAAGTGT 742
QY      235 ProIleGlnProArgHisMetSerValAlAlIeAspLysPheAsnTYrLysLeuProTYr 254
Db      743 GAGGAGCATCCCAAGCATCTGGTGTGGCAGGAAACGACCTGGGTACAGGTTACGTTAC 802
QY      255 SerAlAlIephGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGly 274
Db      803 AGTGGATATTTGGGGGTCTTACTGCCCTTAGCAGAGACGATTTTCAAGGTGAATGA 862
QY      275 PheSerAsnAspPheTyrGlyTyrGlyGlyGluAspAspLeuAlaThrArgThrSer 294
Db      863 TTCTCTAACAACTACTGGGATGGGAGCGAAGACGATGACCTCAGACTCAGGGTTGAG 922
QY      295 MetAlaGlyLeuLysValSerArgTYrProThrGlnIleAlaArgTYrLysMetIleLys 314
Db      923 CTCCAAGAAATGAAATTTCCGGCCCCCTGCCTGAGAGTGGGTAAATATACAAATGCTTTC 982
QY      315 HisSerThrGluAlaThrAsnProValAsnLysCysArgTYrLysIleMetGlyGlnThr 334
Db      983 CACACTAGACAGCAAGCATAGGTGACGACGAGAACGATGAAGCTTTACACCAAGTG 1042
QY      335 LysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTYrLysLeuValAsnLeuGlu 354
Db      1043 TCACGAGTCTGAGAACAGATGGGTGAGTAGTGTCTTATAAATTAGTATCTGTGAA 1102
QY      355 LeuLysProLeuTYrThrArgAlaValValAsp 365
Db      1103 CACAAATCTTATATATATCAACATCACAGTGGAT 1135
```

Search completed: January 9, 2006, 04:25:47
Job time : 260 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 9, 2006, 03:00:11 ; Search time 988 Seconds
(without alignments)
3205.639 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAAVARKSLVLCV.....VDLKEKCRRLRDPFCF 383

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 segs, 413468905 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=nlh
-Q=/cgn2_1/USPTO_spool/US10661430/runat_06012006_123138_6363/app_query.fasta_1.519
-DB=Published_Applications_NA_Main -QMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1
-MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45 -DOCLAM=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10661430 @CCN 1.1.1026 @runat_06012006_123138_6363 -NCPU=6 -ICPU=3
-NO_MAP -LARGEDEBERT -NEG_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	1152	7	US-10-661-430-2
2	718	35.4	2002	3	US-09-925-301-230
3	709	34.9	1839	10	US-11-097-143-18554
4	684	33.7	1853	3	US-09-745-763-84
5	675	33.2	4093	10	US-11-097-143-18553
6	667.5	32.9	3953	5	US-10-087-192-2048
7	666	32.8	1158	8	US-10-466-941-5

8	666	32.8	1158	9	US-10-870-635-5	Sequence 5, Appl1
9	666	32.8	4162	5	US-10-084-817-168	Sequence 168, App
10	663.5	32.7	1116	8	US-10-761-435A-19	Sequence 19, Appl
11	661	32.5	1119	7	US-10-287-226-81	Sequence 81, Appl
12	660.5	32.5	1191	9	US-10-713-970-20	Sequence 20, Appl
13	658	32.4	1116	5	US-10-132-652-33	Sequence 33, Appl
14	658	32.4	1116	10	US-11-105-796-33	Sequence 33, Appl
15	658	32.4	1890	9	US-10-956-157-1114	Sequence 114, App
16	658	32.4	2027	5	US-10-132-652-1	Sequence 1, Appl1
17	658	32.4	2027	10	US-11-105-796-1	Sequence 1, Appl1
18	657.5	32.4	3194	8	US-10-723-860-7737	Sequence 7737, Ap
19	655.5	32.3	6406	9	US-10-713-970-7	Sequence 7, Appl1
20	655	32.3	4080	3	US-09-960-706-487	Sequence 487, App
21	655	32.3	4080	3	US-09-873-319-297	Sequence 297, App
22	646	31.8	1023	5	US-10-132-652-34	Sequence 34, Appl
23	646	31.8	1023	10	US-11-105-796-34	Sequence 34, Appl
24	643.5	31.7	1888	7	US-10-287-226-77	Sequence 77, Appl
25	642.5	31.6	1964	3	US-09-989-722-235	Sequence 235, App
26	642.5	31.6	1964	3	US-09-989-723-235	Sequence 235, App
27	642.5	31.6	1964	3	US-09-989-723-235	Sequence 235, App
28	642.5	31.6	1964	3	US-09-989-279-235	Sequence 235, App
29	642.5	31.6	1964	3	US-09-989-721-235	Sequence 235, App
30	642.5	31.6	1964	3	US-09-989-732-235	Sequence 235, App
31	642.5	31.6	1964	3	US-09-991-073-235	Sequence 235, App
32	642.5	31.6	1964	3	US-09-990-442-235	Sequence 235, App
33	642.5	31.6	1964	3	US-09-991-163-235	Sequence 235, App
34	642.5	31.6	1964	3	US-09-990-604-235	Sequence 235, App
35	642.5	31.6	1964	3	US-09-990-456-235	Sequence 235, App
36	642.5	31.6	1964	3	US-09-989-721-235	Sequence 235, App
37	642.5	31.6	1964	3	US-09-992-598-235	Sequence 235, App
38	642.5	31.6	1964	3	US-09-989-293A-235	Sequence 235, App
39	642.5	31.6	1964	3	US-09-989-735-235	Sequence 235, App
40	642.5	31.6	1964	3	US-09-990-444-235	Sequence 235, App
41	642.5	31.6	1964	3	US-09-991-181-235	Sequence 235, App
42	642.5	31.6	1964	3	US-09-989-730-235	Sequence 235, App
43	642.5	31.6	1964	3	US-09-990-436-235	Sequence 235, App
44	642.5	31.6	1964	3	US-09-993-687-235	Sequence 235, App
45	642.5	31.6	1964	3	US-09-989-724-235	Sequence 235, App

ALIGNMENTS

RESULT 1
US-10-661-430-2
; Sequence 2, Application US/10661430
; Publication No. US20040086995A1
; GENERAL INFORMATION:
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: Kawai, Ziad S.
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS OF USE THEREOF
; FILE REFERENCE: 7148.001
; CURRENT APPLICATION NUMBER: US/10/661.430
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 60/411,242
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-661-430-2

Alignment Scores:
Pred. No.: 4.58e-257 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0
US-10-661-430-1 (1-383) x US-10-661-430-2 (1-1152)

```
QY 1 MetAlaPheArgHisIleuAlaValAlaArgIleuSerIleuValIleuCysAlaVal 20
Db 1 ATGGCTTTTTCGATTTGGAGTCGCCAGACTCAAGTCGTTGCTTACTTGGCCGTT 60
QY 21 LeuIleuValHisIleuMetIleTyTyLysIleProSerIleuTyrgIuansIleuThiIle 40
Db 61 CTTCTATTAAGTTCATGCAATGATTTTAAAGATTCATCGCTTTACGAAACCTTACTATC 120
QY 41 GlySerSerThiIleuIleAlaAspValAspAlaMetGluAlaValIleuGlyAsnThiAla 60
Db 121 GGCTCTCGACCTTATTTGGCGAGTCGACGACGATGAGCAATGAGCAATGCTCGGAAATACGGCT 180
QY 61 SerThiSerAspIleuIleuAspThiTrpAsnSerThiPheSerProIleSerGluVal 80
Db 181 TCCACTTCGGATATCTACTTGAATAGCTGAAATTCACGCTTTTCCACGATTTCTGAAGTT 240
QY 81 AsnGlnThiSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThiLeu 100
Db 241 AATCAGACTAGTTTATGAGAGCAATTCGTCATCTGTTCCCGACCAACAGACTCTT 300
QY 101 GlnPheCysAsnGlnThiProProHisLeuValGlyProIleArgValPheLeuAspGlu 120
Db 301 CAATTCGTATCAGACACCTCCCACTCGTGGACCATCCGATATTCCTCGATGAG 360
QY 121 ProAspPheIleThiLeuGluLysIleTyProAspThiHisIleGlyGlyMet 140
Db 361 CCGGACTTCAAACTCCGAGAAATCTATCCGACACGACCCGGTGGACATGAAATG 420
QY 141 ProLysAspCysValAlaArgHisArgValAlaIleIleValProTyArgAspArgGlu 160
Db 421 CCTAAGGATTTGTGTGCAAGGATCGTGTGCTATTTATTTGGCCCTATGAGATCCGTGA 480
QY 161 AlaHisLeuArgIleMetIleuHisAsnLeuHisSerIleuValLysGlnIleuAsp 180
Db 481 GCACATTTGAGAAATGAATGCTCCCAATTTGCATCGTGTGCTCGCAAAACAACATTTGAC 540
QY 181 TyrAlaIlePheIleValIleGluIleValAlaAsnGlnThiPheAsnArgGlyLysLeuMet 200
Db 541 TATGCAATTTTCAATTTGGAGCAAGTGGCGAATAGACGTTTAATCCGGGAAACTATAG 600
QY 201 AsnValGlyTyArgAspValAlaSerArgLeuTyProTrpGlnCysPheIlePheHisAsp 220
Db 601 AAGCTTGATGATGACGATGATGACGCTCTCAACCATGGAGTCTTCATCTTCAATGAT 660
QY 221 ValAspIleuLeuProGluAspAspArgAsnLeuTyTrpCysProIleGlnProArgHis 240
Db 661 GTCGATTTTACTGCCCGAAGATGACGTTAACCCTGACGTCGTCATTTCAACCATGTCAT 720
QY 241 MetSerValAlaIleAspLysPheAsnTyTrpLysLeuProTySerAlaIlePheGlyGly 260
Db 721 ATGAGTGTAGGATCGATTAATTTCAATTAATCTTCATATTTGGCGGATCTTCGGGCGA 780
QY 261 ILeSerAlaLeuThiLysAspHisIleuLysIleLeuAsnGlyPheSerAsnAspPheTrp 280
Db 781 ATCAGTGACCTAAACAAAAGATCACCTGAAAGAAATCAATGATGATTTTGG 840
QY 281 GlyTyTrpGlyGlyuAspAspAspLeuAlaThiArgThiSerMetAlaGlyLysLeuVal 300
Db 841 GGTGGGGCGGAGGACGACGATTTGGCGAGAACATGATGAGGCTGGACGTAAAGTT 900
QY 301 SerArgTyProTrpGlnIleAlaArgTyLysMetIleLysHisSerThiGluAlaThr 320
Db 901 TCAAGATATCCGACACAAATTTGACAGATATAAATGATTAAGCACTCGACGAAAGCGACG 960
QY 321 AsnProValaLysCysArgTyLysIleMetGlyGlnThiLysArgArgTyTrpThrArg 340
Db 961 AATCCAGTTAATTAATGCCGCTAACAAATTAATGGGCCAAAGCGCGATGAGACAGCT 1020
QY 341 AspGlyLeuSerAsnIleuLysTyTrpLysLeuValAsnLeuGluLysProLeuTyTrpThr 360
Db 1021 GACGGCCTTAAGCAATCTGAAGTAAAGCTCGTAATCTGGAATGAAAGCCTCTCAACCT 1080
```

```
QY 361 ArgAlaValaLysAspIleuGluLysAspCysArgArgGluLeuArgArgAspPhePro 380
Db 1081 CGAGCGTCGTGATTTGCTGTGCAAAAAGACTGCCCGGAGACTCGAAGGACCTTCCA 1140
QY 381 ThrCysPhe 383
Db 1141 ACGTGTTTT 1149

RESULT 2
US-09-925-301-230
; Sequence 230, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 2002
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-230

Alignment Scores:
Pred. No.: 4,14e-83 Length: 2002
Score: 718.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.35% Indels: 2
DB: Gaps: 2

US-10-661-430-1 (1-383) x US-09-925-301-230 (1-2002)
QY 95 ProAspAsnGlnThiTrpIleuGlnPheCysAsnGlnThiProProHisIleuValGlyProIle 114
Db 470 CCAATCTCTCAAGGTTGTGCCCTTACTGTCCAGAAAGCATTCCTCTTAAATGGGTCCTGTG 529
QY 115 ArgVal---PheLeuAspGlnProAspPheLysThiLeuGluLysIleTyProAspThi 133
Db 530 TCGGTGCTCTTATGACCCAGTGCATCATGCGACATGAGATTTGTGAGGCGAATCCCGGTA 589
QY 134 HisAlaGlyGlyHisIleGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 590 GAACCAAGGGGGCGGTACCGCCCTGACAGTTGTGAACCCCGCTCCGAACAGCCATCATTT 649
QY 154 ValProTyArgAspArgGluAlaHisIleuArgIleMetIleuHisAsnLeuHisSerLeu 173
Db 650 GTGCTCATCTGTGCTCCGGAGACACACCTGCGCTGTGCTGCTCAACACCTGCACCCCTTC 709
QY 174 LeuAlaLysGlnIleuAspTyArgAlaIlePheIleValIleGluIleValAlaAsnGlnThi 193
Db 710 TTGACAGCGCAGACAGCTTGCTTATGCACTATGATCATCACACAGCGTGGAAATGGAACA 769
QY 194 PheAsnArgGlyLysLeuMetAsnValGlyTyArgAspValAlaSerArgLeuTyTrpTrp 213
Db 770 TTTTAACGGGCAAAAGCTTTGAACGTTGGGCTCGAAGGCGCTGCGTGAAGAAGTGG 829
QY 214 GlnCysPheIlePheHisAspValaLysPheLeuProGluAspAspArgAsnLeuTyTrpThr 233
Db 830 GACTGCTGTTCTTGACAGATGTGACCTTTGCCGAAATATACCAACATCTGTATGTG 889
QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyTyLysLeu 252
Db 890 TGTGACCCCGGGGAGACCCCGGCATGTTGCCGTTGCTATGAACAAGTTTGGATACAGCTC 949
QY 253 ProTySerAlaIlePheGlyGlyIleSerAlaLeuThiLysAspHisIleuLysIle 272
```


Alignment Scores:

Pred. No.: 6,63e-77 Length: 4093
Score: 675.00 Matches: 161
Percent Similarity: 50.00% Conservative: 59
Best Local Similarity: 36.59% Mismatches: 137
Query Match: 33.23% Indels: 83
DB: 10 Gaps: 11

US-10-661-430-1 (1-383) x US-11-097-143-18553 (1-4093)

```
OY 5 HisLeuAlaValAla-ArgLeuLysSerLeuLeuValLeuCyAlaValLeuLeuVal 24
DB CATATTGCTTGGGAGGTTATCTTATCAATTAATTTTGGAGACGATATAGCATTT-- 2795
OY 24 HisAlaMetIleTyTyLysIleProSerLeuTyTyLysLeuLeuThr----- 1Leu1 41
DB 2794 ----CGGCGCGTGCAAGATATGCTCATATCTTACGGGAAAGCTAGCAGGATGAGCCGG 2739
OY 41 YSerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSe 61
DB 2738 AGGCACTGAA---GCATCCAGGCTGCCGCTTCCCGCTCGCTTATCAAAAGACAGAGA 2682
OY 61 rThSerAspAspLeuAspThrTrpAsnSerThrPhe---SerProIleSerGluVal 80
DB 2681 GCGGAGCAGAGCTCAATGGCGGAGCCAACTCTACATTAAGACATGTGATGGACCGGC 2622
OY 80 lAenGlnThrSerPheMetGluAspIleArgProIleLeu-----Ph 94
DB 2621 AAACCTTACTTCCATCTCCACAAAGACTTTCACGGCTTCCCTGCGGACCAAGAAATTTT 2562
OY 94 eProAspAsnGln-----ThrLeuGlnPheCysAsnGln----- 105
DB 2561 GCGCCCGGACAGAAATCAATCCGCGCTCTTGCCAACTGACATGATCCGATCCCG 2502
OY 106 -----ThrProIleIle 110
DB 2501 TGAATGTAAGCTCGGCTGGCTCTTTCTTTAGACGATTTAAATACGTCGACCTGTTT 2442
OY 110 uValGlyProIleArgValPheLeuAspGluProAspPheLysThrLeuGln--Lys11 129
DB 2441 AGGTGACCCCATCCAGCCCAACAGCAGCAGTGAAGTCACTGACGTTATTTAGGCGGAGCT 2382
OY 129 eTyxProAspThrHisAlaGlyLysIleGlyMetCProLysAspCysValAlaArgHisAr 149
DB 2381 TGAACCTCTTTTGGCGCCCTGGTGGGCGCTTCCAGCTTGAAACCTGAATGCCAGATCA 2322
OY 149 gValAlaIleIleValProTyxArgAspArgGluAlaHisLeuArgIleMetLeuHisAr 169
DB 2321 CGTGGCTATTTGTGGCTTCCGCGATGATACGCCCATCTATTAATTTTCTCGCGCA 2262
OY 169 nLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyxAlaIlePheIleValGlnGlnVa 189
DB 2261 CATCCACCCATTTCCGATGAAACAGCGCATCCCATTCATTCATTTGATGTAGACAGAC 2202
OY 189 lAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyxAspValAlaSerAr 209
DB 2201 CAACGGAAAGCCCTTAATCGGCTGCGCATGATGAACATTTGTTATTTGGAGCCCTTAA 2142
OY 209 gLeuTyxProIleGlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspAr 229
DB 2141 GGTGACAGGTGGAGATGTTTATATATCCACAGATGTCATCTTGGCTTGGAGAGACGG 2082
OY 229 gAsnLeuTyxThrCysPheIleGlnProArgHisMetSerValAlaIleAspLysPheAs 249
DB 2081 CATCTTTTACAACTGCCAGTCCAGCGCGACACATGTCTAGTGGCTATGACACCGCTGAA 2022
OY 249 n-----Ty 250
DB 2021 CTTCAGGCAAGTAGCAAGAGATACATATCATGTCAATGTAATTTGATTCATTTCTTA 1962
OY 250 rLysLeuProTyxSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuTy 270
DB 1962 rLysLeuProTyxSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuTy 270
```

```
DB 1961 TAGGTGCTTATCGATCAATATTTGGAGGTGTTCCGCAATGACCGGTGACCTTCA 1902
OY 270 gLysIleAsnGlyPheSerAsnAspPheTrpGlyTyTrpGlyGlyLysAspAspLeuAl 290
DB 1901 GCGCTTAATGGGTTCTAAACTGTTTGGCTGGGCGGCGGAGATGACAGCATGTC 1842
OY 290 aThrArgThrSerMetAlaGlyLeuLysValSerArgTyxProThrGlnIleAlaArgTy 310
DB 1841 CAACAGGTGAAGCAGCGAAACCTTATCATATCAAGATTCGCGTCACATGCTCGGTA 1782
OY 310 rLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyxLys11 330
DB 1781 CAAGATGCTGAAGCATTCAGAAAGAAAGCCAAATCTTAAGCGGTGATGACCATATAT 1722
OY 330 eMet----- 331
DB 1721 TATTCATATTTAGACTTACAAATATGGAATTTTACCTTCTTTTGGACAGCTATGA 1662
OY 332 -----GlyGlnThrLysArgArgTyxThrArgAspGlyLeuSerAsnLeuTy 347
DB 1661 AAACCTTACAGAAATGCGCATGAGCAAAATAGAA-----CAGATGGAATCACTCGATAA 1608
OY 347 eTyxLysLeuValAsnLeuGlyLeuLysProLeuTyxThrArgAlaValAspLeu 366
DB 1607 GTATTCATCTTACAGCATCAAAATTTCCAACTTTCACCTTGTACTTATGACAGACTA 1550
RESULT 6
US-10-087-192-2048
/ Sequence 2048, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Moritz, David W.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ FILE REFERENCE: 529452000122
/ CURRENT FILING DATE: US/10/087,192
/ PRIOR FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2048
/ LENGTH: 3953
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-10-087-192-2048
Alignment Scores:
Pred. No.: 6.13e-76 Length: 3953
Score: 667.50 Matches: 123
Percent Similarity: 65.06% Conservative: 52
Best Local Similarity: 45.72% Mismatches: 93
Query Match: 32.87% Indels: 1
DB: 5 Gaps: 1
US-10-661-430-1 (1-383) x US-10-087-192-2048 (1-3953)
OY 99 ThrLeuGlnPheCysAsnGlnThrProIleLysLeuValGlyProIleArgVal---Phe 117
DB 551 TCGCTGCCAGCTGGCGCCGAGAGAGTCCCGCTGCTCGTTGGCGCCATGCTGATTCATTT 610
OY 118 LeuAspGluProAspPheLysThrLeuGlnLysIleTyxProAspThrHisAlaGlyGly 137
DB 611 AATATTGCTGATCTGAGCTTTTGGCAAGAAAGAACCCAGATTAAGACGCGGCGGC 670
OY 138 HisGlyMetCProLysAspCysValAlaArgHisArgValAlaIleIleValProTyxArg 157
DB 671 CGTATCTCCCAAGAGACTGTGTCTCTCTCAAGAGTGGCGCATATCATATCCATTCGCT 730
OY 158 AspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGln 177
```

DB 731 AACCGCAGAGATCTCAAAATACGTGCTGATATTATTTGATCCATCCCTCAGCGCAG 790
QY 178 GlnLeuAspTyrAlaIlePheIleValGlnIleValAlaIleGlnIlePheAsnArgGly 197
DB 791 CAACCTGACCTATGCGATCTACGTCATCATCAGCGTGAACACCATGTTCAATCCAGCT 850
QY 198 LysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIle 217
DB 851 AAGCTCTCAATATTGGCTTTCAGAGAGCCCTTGAAGACATATGATTCACATCTTGG 910
QY 218 PheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGln 237
DB 911 TTCAGAGATGAGACCTCATTCGATGAGACGACGTAATGCTACAGAGTGTTCGACAG 970
QY 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIle 257
DB 971 CCACGGCACATTTCTGTTCGCAATGACAGTTGGGTTTACCTGCTGCATATGTTCAAT 1030
QY 258 PheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsn 277
DB 1031 TTGGAGAGTGTCTGTCTTCAGTAACAAACAGTTCTTCCCATTCATGAGATTCCTAAT 1090
QY 278 AspPheTyrGlyTyrGlyGlyGluAspAspArgLeuAlaThrArgThrSerMetAlaGly 297
DB 1091 AATTATTTGGGGTGGGAGAGAGAAAGATGACGACATTTTAAACAGATTAGTTCAATAAGC 1150
QY 298 LeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThr 317
DB 1151 ATCTCATATTCACGTCCTCAAAATGCTGTATGAGGAGGTGTGAATGATCCCGCATTCAMA 1210
QY 318 GluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArg 337
DB 1211 GACAAAGAAAATAGCCCAATCTCTCAGAGGTTTGAACCGGATCCGACATCAAGAGAAAG 1270
QY 338 TrpThrArgAspLysLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysPro 357
DB 1271 ATCGGCTGTGATGTTTGAATCCTACCTACAGAGTGTGATGACAGATACCG 1330
QY 358 LeuTyrThrArgAlaValValAspLeu 366
DB 1331 TTATATATCCCAATCAACAGTGAATC 1357
RESULT 7
US-10-466-941-5
Sequence 5, Application US/10466941
Publication No. US20040214273A1
GENERAL INFORMATION:
APPLICANT: Tatsuji, Seki
APPLICANT: Kazuhito, Fujiyama
TITLE OF INVENTION: Methods for excretory production of glycoproteins
FILE REFERENCE: J198020401
CURRENT APPLICATION NUMBER: US/10/466, 941
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: JP200112519
PRIOR FILING DATE: 2001-01-19
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 1158
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1155)
US-10-466-941-5
Alignment Scores:
Pred. No.: 1.24e-76 Length: 1158
Score: 666.00 Matches: 137
Percent Similarity: 54.69% Conservative: 73
Best Local Similarity: 35.68% Mismatches: 128

Query Match: 32.79% Indels: 46
DB: 8 Gaps: 5
US-10-661-430-1 (1-383) x US-10-466-941-5 (1-1158)
QY 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTyr----- 29
DB 37 CTCGTGCGCGTCTGCGCTCTGACACCTTGCGGTCAACCTGTTTACTACTGCTGCGCCG 96
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
DB 97 GACCTGAGCGCGCTGCCCACTG-----GTGGAGTCTTCACACCGCTG 141
QY 47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
DB 142 CAGGCGGCTCGAACAAGTCCCGCCATCGGGCACTCTCGGGAGAGCTCCGGACCGGA 201
QY 61 -----SerThrSer 63
DB 202 GGGGCGCGCGCGCTCTAGGCGCTCTCCACCGCGCGCGGAGTGGCACTCC 261
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
DB 262 AGCCCACTGTGATTCGCGCTGCGCCCTGAGCACTGACCTGACCTCGGCTC----- 312
QY 84 SerPheMetGluAspIleArgProIleLeuPheAspAsnGlnThrLeuGlnPheCys 103
DB 313 -----CCAGTCCCGCACACACCGCACTGCTGCCGCTGC 351
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
DB 352 CTGAGAGAGTCCCGCTACTAGTGGGCCCCCATGCTAGTGAATGATTCAGTTCGCTGAG 411
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyHisIleGlyMetProLys 142
DB 412 CTGGAGCTGTGCGCAAGAGAACCCAAATGTGAAGATGGCGCGCTATGCCCGCCAGG 471
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
DB 472 GACTGCTCTCTCTCTCAAGAGGCGCATCATATTCATTCGCAACCGGACGAGAGCAC 511
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnIleLeuAspTyrAla 182
DB 532 CTCAGTACTGCTATATATTATTGACACCCAGCTCTCAGCGCAGCAGCTGACTATGCG 591
QY 183 IlePheIleValGlnIleValAlaAsnGlnThrPheAspArgGlyLysLeuMetAsnVal 202
DB 592 ATCTATGTTATCAACACGCGGAGACACTATATTCATATGCTGAAGCTCCCAATGTT 651
QY 203 GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAsp 222
DB 652 GGCCTTCAGAGACCTTTGAAGACTATGACTACCTGCTTGTGTACTGACGTGAC 711
QY 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
DB 712 CTCATTCAATGATGATCCATATATGCTACAGGTGTTTTCACAGCAGCGCACATTTCC 771
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyLysIleSer 262
DB 772 GTTGCAATGATTAAGTTTGATTCACGCTTACCTTATGTTAGTATTTTGGAGGTCTCT 821
QY 263 AlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTyrGlyTyr 282
DB 832 GCTCTAAGTAACAAACAGTTCTTAACCATCAATGATTCCTTAATAATATTGGGCGTGG 891
QY 283 GlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
DB 892 GGAAGAGAGATGATGATTTTAAACAGATTATTTAGAGGCACTGCTATATCTCGC 951
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
DB 952 CCAATGCTGTGTGCGAGGTGTGCTGATGTCGCGCATCAAGAGACAAAGAAATGAA 1011

```
QY 323 ValAaenLysCyAargTyrLysIleMetGlyInThrLysArgArgTrpThrArgAspGly 342
    |||:::
    1012 CCCAATCTCCAGAGGTTTACCGAATTCGACACACAAAGAGACATGCTCTGATGCT 1071
QY 343 LeuSerAsnLeuLysTyrLysLeuValAsnLeuGlyLeuLysProLeuTyrThrArgAla 362
    |||:::
    1072 TTGAATCTCACTCACCTACCAAGGTGCTGATGTACAGAGATACCATTTGATATCCCAATTC 1131
QY 363 ValValAaPLeu 366
    |||:::
    1132 ACAGTGACATC 1143
Db 1132 ACAGTGACATC 1143

RESULT 8
US-10-870-635-5
; Sequence 5, Application US/10870635
; Publication No. US20050143564A1
; GENERAL INFORMATION:
; APPLICANT: Tatsuji, Seki and Kazuhiko, Fujiyama
; TITLE OF INVENTION: Methods for manufacturing of glycoproteins
; TITLE OF INVENTION: Having human-type glycosylation
; FILE REFERENCE: 62, 736B
; CURRENT APPLICATION NUMBER: US/10/870,635
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 09/857,651
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
US-10-870-635-5

Alignment Scores:
Pred. No.: 1.24e-76 Length: 1158
Score: 666.00 Matches: 137
Percent Similarity: 54.69% Conservative: 73
Best Local Similarity: 35.68% Mismatches: 128
Query Match: 32.79% Indels: 46
DB: 9 Gaps: 5

US-10-661-430-1 (1-383) x US-10-870-635-5 (1-1158)
QY 14 LeuLeuValLeuCyAlaValLeuLeuValHisAlaMetIleTyr----- 29
    |||:::
    37 CTGCTGGCCGCTGCGCTGCGCTGACCTTGCGCTTACCTGCGTGGCGGC 96
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrIleuIle 46
    |||:::
    97 GACCTGAGCGCGCTGCCCAACTG-----GTGCGAGTCTCCACACCGCTG 141
QY 47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
    |||:::
    142 CAGGCGCGCTCGAACAAGTCCGCCCATCGGCACTCTCCGGGAGAGCTCCGGACCGGA 201
QY 61 -----SerThrSer 63
    |||
    202 GGGGCGCGCGCGCGCTCTCTAGGCGCGCTCTCCCAAGCGCGCGCGGGTGCGCACTCC 261
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
    |||:::
    262 AGCCCAAGTCGTGATTTCTGGCCCTGGCCCGCTAGGAACTTAACTTCGCTC----- 312
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
    |||:::
    313 -----CCAGTGGCCCAACACACCGCACTGCTGCTGCCCGCTGC 351
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro---Asp 122
    |||:::
    352 CCTGAGAGTCCCGCTACTAGTGGGCCCATGCTGATTTGAAGTTTAAACATGCTGTGGAC 411
```

```
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyGlyHisGlyMetProLys 142
    |||:::
    412 CTGGAGCTCTGGCCAAAGCAGAACCCAAATGTGAAGATGGGCGCCGCTATCCCTCAGG 471
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
    |||:::
    472 GACTCGTCTCTCCACAGAGTGGCCATTCATTCATTCATTCCTCGCAACCGGACGAGAGAC 531
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrAla 182
    |||:::
    532 CTCAGAGTACGTGCTATATTTATTTGACCCAGCTCGACCGCCGACAGACCTGGACATATGCG 591
QY 183 IlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetSerVal 202
    |||:::
    592 ATCTATGTATATCAACAGCGGAGACACTATATTAATCGTGCTAGCTCAATATGTT 651
QY 203 GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAsp 222
    |||:::
    652 GGCCTTCAAGAAAGCTTGAAGAGCTAGTACACTGCTTTGTGTAGTACGTTGAC 711
QY 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
    |||:::
    712 CTCATTCCAATGAATGAACCATATATCGTACAGGTGTTTTCACAGCCACGCGCACTTTCC 771
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyGlyLysSer 262
    |||:::
    772 GTTGCATATGAGATTAATTTTGATTCAGCTTACCTTATGTTCACTATTTTGGAGGTGCTCT 831
QY 263 AlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspPheTyrPglTyr 282
    |||:::
    832 GCTTAAGTAAACAACAGTTTCTAACCATTCATGATGATTTCCATATATATTTGGGGCTGG 891
QY 283 GlyIleGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
    |||:::
    892 GGAGGAGAAAGATGATGACATTTTAAACAGATTGATTTTAAAGGATGTCATATATCTCGC 951
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrArgPro 322
    |||:::
    952 CCAAATGCTGCTGGTGGAGGTGCTGCATGATCCGCCATCGCAAGACAAAGAAATGAA 1011
QY 323 ValAaenLysCyAargTyrLysIleMetGlyInThrLysArgArgTrpThrArgAspGly 342
    |||:::
    1012 CCCAATCTCCAGAGGTTTACCGAATTCGACACACAAAGAGACATGCTCTGATGCT 1071
QY 343 LeuSerAsnLeuLysTyrLysLeuValAsnLeuGlyLeuLysProLeuTyrThrArgAla 362
    |||:::
    1072 TTGAATCTCACTCACCTACCAAGGTGCTGATGTACAGAGATACCATTTGATATCCCAATTC 1131
QY 363 ValValAaPLeu 366
    |||:::
    1132 ACAGTGACATC 1143
Db 1132 ACAGTGACATC 1143

RESULT 9
US-10-084-817-168
; Sequence 168, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Pilon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; PRIOR APPLICATION NUMBER: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 168
; LENGTH: 4162
; TYPE: DNA
```


Db 160 AAAATAGACATTGGAGCGTTGGTCTAGCTGAGAAATAGATCATCTTAATAATTAGA 219
Qy 68 AspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPheMetGlu 87
Db 220 GAC-----TCAGTCATCATATTGGTAGTCTGTGGAG 252
Qy 88 AspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPhe----- 102
Db 253 GAT-----GATCGAAAAGATTCAAAAGCAATTTCAGCCAGAGTGTGGC 297
Qy 103 -----CyAsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAsp 119
Db 298 TCACCCGCGCTGCTGAGAGATGCCCGCTGTGGGCCCATGCTGATGATTAAAC 357
Qy 120 GluPro---AspPheSerThrLeuGluValIleTyrProAspThrHisAlaGlyGlnHis 138
Db 358 ATGCTGTGGAGCTGAGCTGCGGCAAGCAGAACCCAAAGTGAAGTGGAGCGGCCGC 417
Qy 139 GlyMetProLysAspCysValAlaArgHisArgValAlaIleValProTyrArgAsp 158
Db 418 TATGCCCGCCAGGAGCTGCGCTCTCTCTCAAGAGTGGCCATCATCTTCATCCGCAC 477
Qy 159 ArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnGln 178
Db 478 CGCGAGAGACACCTCAAGTACTGCTATATTATTTGCAACCCAGTCTCGACGCGCAG 537
Qy 179 LeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLys 198
Db 538 CTGAGCTATGAGCATATGTATATCAACAGCGCGGAGACACTATATTCATATGTTGCTAAG 597
Qy 199 LeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePhe 218
Db 598 CTCCTCAAGTGGCTGCTTCAAGAGAGCTTGAAGAGACTAGATCAACCTGCTTGTGTT 657
Qy 219 HisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnPro 238
Db 658 AGTGAAGTGAAGCTCATCTCAATGAATGACCAATGCGTACAGGCTTTTTCACAGCCA 717
Qy 239 ArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePhe 258
Db 718 CGGCACTATTCGGTGAATGATGAATGATTGATTCAGCTTATGCTTATGATATTTT 777
Qy 259 GlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAsp 278
Db 778 GGAAGGCTCTGCTCTAGTAAACAACAGTTTCTAACATCAATGATTTCTTAATAT 837
Qy 279 PheTyrGlyTyrGlyGluAspAspLeuAlaThrArgThrSerMetAlaGlyLeu 298
Db 838 TATGGGCGCTGGGAGAGAAATGATGACATTTTAAACAGATTAGTTTGAAGGCGATG 897
Qy 299 LysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGlu 318
Db 898 TCTAATCTCGCCCAATGCTGTGCGGAGGTGCGATATCGCCATCAAGAGAC 957
Qy 319 AlaThrAspProValAsnLysCysArgTyrLysIleMetGlyIleThrLysArgArgTyr 338
Db 958 AAAAAAAGAACCCATCTCAGAGGTTTGACCCGAATTCACACACAAAGAGACAATG 1017
Qy 339 ThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeu 358
Db 1018 CTCTCTGATGATTTGAACTCATCTCAAGGCTGCTGATGATACAGATACCATG 1077
Qy 359 TyrThrArgAlaValAspLeu 366
Db 1078 TATACCCCAATCAACAGTGAATC 1101

APPLICANT: Berghs, Constance,
APPLICANT: Boldos, Ference,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eissen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangolli, Esba A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: MaYankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Patutajan, Weera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vernet, Corine A.M.,
APPLICANT: Zernusen, Bryan D.,
APPLICANT: Zhong, Wei
FILE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 81
LENGTH: 1119
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1116)
US-10-287-226-81
Alignment Scores:
Pred. No.: 5,35e-76 Length: 1119

Score:	661.00	Matches:	141
Percent Similarity:	56.87%	Conservative:	70
Best Local Similarity:	38.01%	Mismatches:	133
Query Match:	32.55%	Indels:	27
DB:	7	Gaps:	7

US-10-661-430-1 (1-383) x US-10-287-226-81 (1-1119)

OY	7	AlaValAlaArgLeu-LysSerLeuValLeuCysAlaValLeuLeuValHisAla	26
Db	24	GCTGAGGCGCGCTGCGAGAGCTGTGCTCTCTCTGCGCTGCACTTCTCTGCGCCG	83
OY	26	AmetIleTyrIleProSerLeuTyrGluAsnLeuThrIleGlySerSer--ThrIle	45
Db	84	CATCCTTAC--TTTGACGTCTACGCCACAGACCTGAGCCTTCTTACGCGCTTACGTG	140
OY	45	uIlAlaAspValAspAlaMetGluValValLeuGluAlaSerThrAlaSerThrSerAsp	65
Db	141	CCGAGGCCCTGCCATGCTCCCTCACCCAGCTGTACACACAGACAGACGACGACAACTG	200
OY	65	pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPh	85
Db	201	CTCCGGCGCCCAACGCCACCGCTCTAGCTCTCGGAGCTCCCTAGAGTCCCAAGTCC-----	255
OY	85	eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnG	105
Db	256	-----CTGCCCGGTCCCAACGCGCTCCACCGCTCCACCTGTGCTCTGA	296
OY	105	nThrProProHisLeuValGlyProIleArgValPheLeuAspGluProAspPheLysTh	125
Db	297	CTCGGCACCTGGTCTTGTGGAGACTGTGACTGTGACTGATTCACCTCAACC-----ATGCC	350
OY	125	rLeuGluLysIleTyr-----ProAspThrHisAlaGlyIleAspGlyMetProLys	142
Db	351	CCTGAGCGGGGTCACACAGGAGAACCCAGCGCTGCATATGGGGCGCGCATACACACCGCC	410
OY	142	aAspCysValAlaIaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaH	162
Db	411	CGACTGCACCCCAAGCCACAGCGGTGGCGGTCATCATCCCTTTAGACACCGGGAAACACA	470
OY	162	sLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnGlnLeuAspTyrAl	182
Db	471	CTTGCGGTACTGGCTCATTATTAACACCCCATCTTGAGCGGACAGCGGCTGGCGCTACGG	530
OY	182	aIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnVa	202
Db	531	CGCTCATGTCAATCAACACAGATGTGTAGAGACACTTCAACCGGGCCACAGCTGTAACT	590
OY	202	IglTyr-----AspValAlaSerArgLeuTyrProTyrGlnCysPh	216
Db	591	GGGCTTCTCTAGAGCGCGCTGAGAGAGATGCGCC-----TATGACTGCTT	635
OY	216	eIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProI	236
Db	636	CATCTTCAAGCATGTGACCTGTGCTCCCATGATGACCGCACTTATACCGCTGCGGCGA	695
OY	236	eGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAl	256
Db	696	CCAAACCCGCGCACTTGGCATTTGCCATTCSCATGACCAAGTATGGCTTCGATGCTGG	755
OY	256	aIlePheGlyGlyIleSerAlaLeuThrLysAspHisLysLeuLysAlaIeAsnGlyPheSe	276
Db	756	CTACTTTGAGAGGTGTCTCAGCGCTGTGATGAGGCTCAAGTTCTTGAGATCATAGGCTTCCC	815
OY	276	rAsnAspPheTrpGlyTyrProGlyValGluAspAspArgLeuAlaThrArgHisSerMetAl	296
Db	816	CATATGATCTGGGGCTGGGGGTGGCGAGATGATGACATTTCAACCGCATCTTCCTTGAC	875
OY	296	aGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSe	316
Db	876	TGGAGATGAAGATCTCAGCCCAACATCCGAATTTGGCGGCTGACCGCATGATCAAGACAGA	935
OY	316	rThrGluAlaThrAsnProValAsnLysCysValArgTyrLysIleMetGlyGlnThrLysAr	336

RESULT 12

US-10-713-970-20
; Sequence 20, Application US/10713970

; Publication No. US20050106664A1
; GENERAL INFORMATION:
;

APPLICANT: Contreras, Roland
; APPLICANT: Callewaert, Nico L. M

APPLICANT: Vervecken, wouter
APPLICANT: Kaigorodov, Vladimir
;

```

; TITLE OF INVENTION: MODIFICATION OF PROTEIN
; FILE REFERENCE: 17106
;

```

; CURRENT APPLICATION NUMBER: US/10/7
 ; CURRENT FILING DATE: 2003-11-14
 ;

```

; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2.1

```

; SEQ ID NO 20
;
; LENGTH: 1191

END DDD

```

; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-10-113-970-20

Alignment scores:
Pred. No.:

Percent Similarity

Best Local Similarity
Query Match:

DB: 10 001 426 1

OS-10-661-430-1 (1-383) X OS-10-713-970-20 (1-1131)

```

00 /Alldat1dat2u1u2u3u4u5u6u7u8u9uaubucudueufu10fu11fu12fu13fu14fu15fu16fu17fu18fu19fu20fu21fu22fu23fu24fu25fu26fu27fu28fu29fu30fu31fu32fu33fu34fu35fu36fu37fu38fu39fu40fu41fu42fu43fu44fu45fu46fu47fu48fu49fu50fu51fu52fu53fu54fu55fu56fu57fu58fu59fu60fu61fu62fu63fu64fu65fu66fu67fu68fu69fu70fu71fu72fu73fu74fu75fu76fu77fu78fu79fu80fu81fu82fu83fu84fu85fu86fu87fu88fu89fu90fu91fu92fu93fu94fu95fu96fu97fu98fu99
01 .....
02 .....
03 .....
04 .....
05 .....
06 .....
07 .....
08 .....
09 .....
10 .....

```

[illegible]

שם	מספר	תאריך	מספר	תאריך
1.00	1.00	1.00	1.00	1.00
2.00	2.00	2.00	2.00	2.00
3.00	3.00	3.00	3.00	3.00
4.00	4.00	4.00	4.00	4.00
5.00	5.00	5.00	5.00	5.00
6.00	6.00	6.00	6.00	6.00
7.00	7.00	7.00	7.00	7.00
8.00	8.00	8.00	8.00	8.00
9.00	9.00	9.00	9.00	9.00
10.00	10.00	10.00	10.00	10.00
11.00	11.00	11.00	11.00	11.00
12.00	12.00	12.00	12.00	12.00
13.00	13.00	13.00	13.00	13.00
14.00	14.00	14.00	14.00	14.00
15.00	15.00	15.00	15.00	15.00

[illegible][illegible]

-----60

[illegible]

```
D0      ZI7ICCCCCGGGAGCCLCCCGACCCGGAGGCCCCCCTCCATCAGTGCCTTC  2
```

```
Ov     61 -----SerThrSerAsnAsnLeuLeuAsnThrTrpAsnSerThrPheSer   75
```

[illegible]

```

26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1
```

334	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	365
335	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
336	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
337	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
338	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
339	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
340	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
341	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
342	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
343	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
344	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
345	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
346	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
347	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
348	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
349	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
350	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
351	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
352	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
353	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
354	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
355	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
356	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
357	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
358	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
359	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
360	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
361	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
362	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
363	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
364	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	
365	AAAAAAAAAAAAAAAAAAAA	CCCCCCCCCCCCCCCC	

```
06 Acnabncg]ntHrtaug]npbhc'vabancg]ntHrBrorDroH[e]aienVa[clvBroIeArx 11
```

[illegible]

116 VajPhelauaengInpro---denPhelvetThriangCilulvajiLeTvrProaSnThrhIs 133

[illegible][illegible]


```

QY 135 AAGAGYGLNIEGLYMERPROLYSAPRCYSVALAIAARGHIEARGVALAIAIEILEV1 154
    |||||
DB 484 AAGGGGCGCGCGCTAGGCCCGCCAGGACCTCGCTCTCCCTCACAGGTGCCATCATATT 543
QY 155 PROTYRARGAPRARGIUAIAHISLEUARGIILEMELEUHIASLEUHIISERLEU1 174
    |||||
DB 544 CCAATTCGCCGACCGGACGAGACCTCAAGTACCTGCTATATTTATTTGACCCAGCCCTCG 603
QY 175 AIAIYSGINGINLEUASPRYRZALAIIEPHEILEVALIGUINVALAIAISNGINHPRE 194
    |||||
DB 604 CAGCGCGACGACCTGACATGACCTATGATTAATCAACAGCGGACGACACCTATATTC 663
QY 195 AENATRGYLYLEUWERASNAVALIGIYRARGVALAIAISERARGLEUYYRPROTGIN 214
    |||||
DB 664 AATCGGCTAAGCTCTCATATGTGGCTTTCAAGAAAGCTTTGAAGACATATACATACACC 723
QY 215 CYPHEILEPHEHISAPRYALAPRLEUENPROGLIUSAPRARGANLEUYYRTHRCYS 234
    |||||
DB 724 TCGTTGTGTAGTGAAGCTGACCTCATTTCCAAATGAAATGACCAATGACCTGACAGTGT 783
QY 235 PROILEGINPROARGHISMETSERVALAIAIEASPRYPHEASNTYRILYSLEUPROTYR 254
    |||||
DB 784 TTTTCACAGCCAGCGCACATTTCCGTTGCAATGATGATGATTTGATTCAGCTACCTTAT 843
QY 255 SERAIAIEPHEGIGYLYLIESERAIALAEUTHRYLASPHISLEUULYSYLIEANGLY 274
    |||||
DB 844 GTTCAGATATTTGAGAGGTCTCTGCTCTAAGTAAACAACAGTTTTCACATCAATGGA 903
QY 275 PHESEASNAAPRPHETRPGLYTRPGIYGIUASPRAPRLEUALATHRARGTHSER 294
    |||||
DB 904 TTTCCAAATATATTTAGGGGCTGGGGGAGGAAGATGATGACATTTTAAACAGATTAGTT 963
QY 295 METALAGLYLEULYVALSERARGTYRPROTHGINLIEALARGTYRILYSMETILYLS 314
    |||||
DB 964 TTTAGAGGATGTCATATCTCGCCCAATGCTGTGTCGGGAGGTGTCATGATGATCCCC 1023
QY 315 HISSETRGUALATHRANPROVALAENLYSCYARGTYRILYSLEUGLYNTHR 334
    |||||
DB 1024 CACTCAAGAGACAAAATAATGAAACCAATCTCCAGAGGTTTGACCGAATTCACACACA 1083
QY 335 LYASARGTYRTPRTHRARGAPRGLYLEUSERASMLEULYSYRILYSLEUVALASNGIN 354
    |||||
DB 1084 AAGAGACAAATGCTCTGATGGTTGAATCACTACCTACAGGTGCTGATGATGACG 1143
QY 355 LEULYSPROLEUTYRTHRARGALAVAIYALAPRLEU 366
    |||||
DB 1144 AGATACCATTTGATATCCCAATACACAGTGCATC 1179

RESULT 13
US-10-132-652-33
; Sequence 33, Application US/10132652
; Publication No. US20030013173A1
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-acetyl-D-glucosamine
; FILE OF INVENTION: Beta-1,4-Galactosyltransferase, betaGal-T2
; CURRENT APPLICATION NUMBER: US/10/132,652
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US/09/118,464A
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652-33

Alignment Scores: 1.33e-75 length: 1116
Pred. No.:

```

```

Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
DB: 5 Gaps: 6

US-10-661-430-1 (1-383) x US-10-132-652-33 (1-1116)
QY 7 AIAVALAARGLEU-LYSELEULEUVALLEUCYSVALAIALEULEULEUVAIHIAI 26
    |||||
DB 24 GCTGAGGCGGCTCTCAAGGCTGTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83
QY 26 AMETILEYRILYSILEPROSERLEUTYRGLUASLEUTHRIIEGLYSERSE--THILE 45
    |||||
DB 84 CATCTCTAC--TTGAAGTACAGCCAGACCTGGCGCTTTCAGCGGCTTCAGTGC 140
QY 45 UILEALASPRVALAPRALAMETGLUALAILEUGLYANETHRALASERTHISERAPAS 65
    |||||
DB 141 CCGAGGCGCGCTGCCATGCTCCACCCAGCTGTCAGACGACGACGACGACGACGACGACG 200
QY 65 PLEULEUASPRTHRTPRANSETRHSPHESEPROLIESEERGIUVAIAENGINTHSERPH 85
    |||||
DB 201 CTCCGCGCCCAAGCCACCGCTTACGTCGGGCTCCCTGAGTCCCAAGTCC-- 255
QY 85 EMEGLIUSAPRIEARPROLIELEUPHESPROASNGINTHLEUGINPHECYASNGI 105
    |||||
DB 256 -----CTGCCGCTCCACCGCTCCACCGCTCCACCGCTCCGCTGCTGA 296
QY 105 NTHRPROBHIISLEUVALIGIYPROLIEARGVALAPHELEUASPRGIPRO--ASPRHELY 124
    |||||
DB 297 CTGCGCACCTGGCTCTGTTGGGACAGCTGTCATGACGTTTCACTTCACTTCACTTCACTT 356
QY 124 ETHRIEUGIUSYILETYRPROASPRTHRHIAIAGLYGIHISGLYMERPROLYSAPRCY 144
    |||||
DB 357 GCGGGGACAGAGGAAACCCAGCGCTGCTATGAGCGCGCGCATACACACCGCGGACTG 416
QY 144 EVALIAARGHISARGVALAIAIELEVALIYPROTYRARGAPRARGIUAIAHISLEUAR 164
    |||||
DB 417 CACCCGACCGGACGAGGTGGGGTGCATATCCCTTTAGACACGGGAAACACACCTGGG 476
QY 164 GILEMELEUHIASLEUHIISERLEUULYSGINGINLEUASPRYRZALAIIEPH 184
    |||||
DB 477 CTAAGGCTCCCATCTTCAACCCCATCTTGAGGGGACGCGGCTGCGCTACGCGGTCA 536
QY 184 ELIEVALIGUINVALAIAENGINTHRPHESAPRARGIYLYSLEUASNAVALIGIYTY 204
    |||||
DB 537 TGTCAACAACGATGCTGAGACACCTTCAACCGGCGCAAGCTGTAAAGTGGGCTT 596
QY 204 r-----ASPRVALASERARGLEUTYRPROTGPINCYEPHEIIEPH 218
    |||||
DB 597 CCTAAGCGCGTGAAGAGATGCCGCC-----TATGACTGCTTATCTT 641
QY 218 EHIISAPRYALAPRLEUENPROGLIUSAPRARGANLEUYYRTHRCYEPROILEGINPR 238
    |||||
DB 642 CAGCGATGAGACCTGATCCCATGATGACCGCAACCTATACCTGCGGCGGACCAACC 701
QY 238 CARHISMETSERVALAIAIEASPRYPHEASNTYRILYSLEUPROTYRSEALAIIEPH 258
    |||||
DB 702 CCGCACTTGTCCATTTGCCATGACGACAGTTGGCTTCCGGCTCTCTATGCTGCTACTT 761
QY 258 eGLIYLIIESERAIALAEUTHRYLASPHISLEUULYSYLIEANGLYPHESEASNAAS 278
    |||||
DB 762 TGGAGGTGTGTCAAGGCTGAGTAAGGCTCAAGTTTCTGAGAAATCAATGCTTCCCAITGA 821
QY 278 PRHETRPGLYTRPGIYGIUASPRAPRLEUALATHRARGTHRSEMERALIGIYLE 298
    |||||
DB 822 GTACTGGGCTGGGCTGGGAGATGATGACATCTTCAACCGGAGATCCCTGATGCGGAT 881
QY 298 ULYSVALSERARGTYRPROTHGINLIEALARGTYRILYSMETILYLSHISERTHNGI 318
    |||||
DB 882 GAAAGATCTACGCGCAACATCCGGAATCGCGGCTTACGACATGACGACGACGACGACG 941
QY 318 UALATHRANPROVALAENLYSCYARGTYRILYSLEUGLYNTHRYARGARGTYR 338

```

```
Db      942 CAAGCATTAAGCACTTAACCTTCAGAGGTTTACCAAGATTCAAAACACGAACTGACCAT 1001
Qy      338 pthraagaagaglyleuSerSerneuLyrylySleuValAsnLeuGluLeuLySprole 358
Db      1002 GAAGCGGAGCGGATGGGTGAGTGGGTACAGAGCTTGTGAGAGTGTCTCGGCAACCACT 1061
Qy      358 uTyTrArGaAlaValAlaAspleu 366
Db      1062 CTTCAACCAATATACAGAGTGGACATT 1086

RESULT 14
US-11-105-796-33
; Sequence 33, Application US/11105796
; Publication No. US20050181437A1
; GENERAL INFORMATION:
; APPLICANT: CLAUSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; FILE REFERENCE: 04305/100521-US1
; CURRENT APPLICATION NUMBER: US/11/105,796
; PRIOR FILING DATE: 2005-04-13
; PRIOR APPLICATION NUMBER: 09/118,464
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-105-796-33
```

Alignment Scores:

```
Prod. No.: 1.33e-75 Length: 1116
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
Gaps: 6
```

US-10-661-430-1 (1-383) x US-11-105-796-33 (1-1116)

```
Qy      7 AlavAlaArghleu-LySerSerleuValleuCyAlaValleuLeuValHisAl 26
Db      24 GCTGAGCGGCTGCGAAGCTGTGCTCTTCTGTGCTGTGACATTCCTCGGGCGCT 83
Qy      26 aMetIleTyrlYsIleProSerleuTyrgluAsnLeuThrIleGlySerSer---ThrIe 45
Db      84 CATCTCTTAC---TTTGACGCTTACGCCACGACCTGCGCTTCTTCCAGCCGCTTCACTGC 140
Qy      45 uIleAlaAspValaAspAlaMetGluAlaValleuGlyAsnThrIleAspThrIleSerAspAs 65
Db      141 CCGAGCGCTGCGCCATCCCTCCACCCAGCTGCTGACGAGCGAGCGAGCGAGCAACCATG 200
Qy      65 pLeuLeuAspThrIleTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPh 85
Db      201 CTCCCGGCGCAACGCCACCGCTCTAGCTCCGGGCTCCCTGAGGTCCCAAGTGC----- 255
Qy      85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrleuGlnPheCyAsnGln 105
Db      256 -----CTGCCCGGTCACACGCTGCCACGCTGCCACGCTGCTCTCTGA 296
Qy      105 nThrProGHisleuValGlyProIleArgValaPheLeuAspGluPro---AspPheLy 124
Db      297 CTCGCCACCTGTGCTTGTGGGACAGCTGATGAGTTCACTTCCACCCATGCCCTGGA 356
Qy      124 sThrleuGluIuYsIleTyProAspThrHisAlaGlyValHisGlyMetProLyAspCy 144
Db      357 GCGGCTGCAAGGAGAAACAGGCGCTGCTAGTGGCGCGCGGAGACACCGCCCATG 416
Qy      144 sValAlaArgHisArgValaAlaIleIleValaProTyTrArGaAspArgGluAlaHisleuAr 164
```

```
Db      417 CACCCCAAGCCAGACGGTGGCGGTATCATCCCTTTAGACACCGGGAAACACCACCTGGC 476
Qy      164 gIleMetLeuHisAsnLeuHisSerleuLeuAlaIuSgInLeuAspTyTrAlaIlePh 184
Db      477 CTACTGGCTTCCATATCTACACCCCATCTTGAGCGCGGACCGGCTCGGCTACGCGCTTA 536
Qy      184 eIleValGluGlnValaIleAsnGlnThrPheAsnArgIlyLySleuMetAsnValGlyTy 204
Db      537 TGCATCAACACGATGATGAGACACCTTCAACCGGCGCAAGCTGCTTACCTGGGCTT 596
Qy      204 r-----AspValaIleSerArgleuTyrrProIleGlnCyPheIlePh 218
Db      597 CTTAGAGCGCTGAGACGAGATGCCGCC-----TATGACTGCTTCACTT 641
Qy      218 eHisAspValaAspLeuLeuProGluAspAspArgAsnLeuTyTrhCySproIleGlnPr 218
Db      642 CAGCGATGTGACCTGTGTCCCATGTGATGACCGCAACCTATACCGGTGGCGGACCAACC 701
Qy      238 oArghIleMetSerValaIleAspIleAspIleAspIleAspIleAspIleAspIleAsp 258
Db      702 CCGCAGCTTTGCCATTTGCGATGGAACAAGTTGGCTTCCGGCTTCCCTATGCTGGCTACTT 761
Qy      258 eGlyGlyIleSerAlaLeuThrIleAspIleAspIleAspIleAspIleAspIleAspIle 278
Db      762 TGGAGGTGTGTACAGGCTGTAGTAAGGCTGAGTTTGAATAATCATGCTTCCCAATG 821
Qy      278 pHetTrpGlyTyTrpGlyGlyIuAspAspAspLeuAlaThrArgThrSerMetAlaGlyLe 298
Db      822 GTACTGGGCTGGGCTGGGCTGGAGATGATGATCATCTTCAACCGGATCTCCCTGACTGGAT 881
Qy      298 uLySValSerArgTyrrProThrGlnIleAlaArgIlySleuMetIleLyHisSerThrGl 318
Db      882 GAAGATCTACAGCCCGCAGACATCCGAATCGCGCTTACCGCATGATCAAGACACACCGCA 941
Qy      318 uAlaThrAspProValaAspIleAspIleAspIleAspIleAspIleAspIleAspIle 338
Db      942 CAAGCATTAAGCACTTAACCTTCAGAGGTTTACCAAGATTCAAAACACGAACTGACCAT 1001
Qy      338 pthraagaagaglyleuSerSerneuLyrylySleuValAsnLeuGluLeuLySprole 358
Db      1002 GAAGCGGAGCGGATGGGTGAGTGGGTACAGAGCTTGTGAGAGTGTCTCGGCAACCACT 1061
Qy      358 uTyTrArGaAlaValAlaAspleu 366
Db      1062 CTTCAACCAATATACAGAGTGGACATT 1086
```

RESULT 15

```
US-10-956-157-1114
; Sequence 1114, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Myeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1114
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1114
```

Alignment Scores:

```
Pred. No.: 3.21e-75 Length: 1890
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
Gaps: 6
```

US-10-661-430-1 (1-383) x US-10-956-157-1114 (1-1890)

```
Oy 7 AlaValAlaArgLeu-LysSerLeuLeuValLeuCysAlaValLeuLeuValAlaHisAl 26
Db 76 GCTGGAGGCGCTCTCAAGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 135
Oy 26 AmetIleTyrIleProSerLeuTyrGluAsnLeuThrIleGlySerSer---ThrIle 45
Db 136 CATCCTCTTC---TTTGAGCTCTAGCGCCAGACCTGGCGCTTCTTCAAGCGCTTCACTGC 192
Oy 45 uIleAlaAspValaAspAlaMetGluAlaValLeuGlnAsnThrAlaSerThrSerAspAs 65
Db 193 CCGAGAGCCCTGCCATGCCCTCCACCCAGCTGTAGACAGACAGACAGACAGACAGCACTG 252
Oy 65 pleuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPh 85
Db 253 CTCCCGGCCCAAGCCACCGCTCTAGCTCCGGGCTCCCTGAGGTCCCAAGTCC--- 307
Oy 85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGln 105
Db 308 -----CTGCCGGTCTCCACGGCTCCACGCTGCCACCTGTCTGA 348
Oy 105 nThrProHisLeuValGlyProIleArgValPheLeuAspGluPro---AspPheIle 124
Db 349 CTGGCACCTGGTCTTGTGGCAGACTGCTGATCGAGTTCACTCACCCATGCCATGCCCTGGA 408
Oy 124 sThrLeuGluIleTyrProAspThrHisAlaGlyIleGlyHisGlyMetProIleAspCy 144
Db 409 GCGGGTGCAGAGGAGAACCCAGCGGTCTCATGGCGGCCGATACACACCCGCCGACTG 468
Oy 144 sValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHisLeuArg 164
Db 469 CACCCAGCCCGAGAGCGTGGCGTCTCATCCCTTTAGACACCGGAGACACCACTGCG 528
Oy 164 gIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnGlnLeuAspTyrAlaIlePh 184
Db 529 CTACTGGGCTCCACTATCTACACCCCATCTTGAAGGGCGAGCGGCTGCGCTACGGGGCTCA 588
Oy 184 eIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyIleLeuMetAsnValGlyTyr 204
Db 589 TGTCTATCAACGACATGTGTAGAGACCTTCAACCGGCCAAGCTGTCTTAACGTGGGCTT 648
Oy 204 r-----AspValAlaSerArgLeuTyrProTyrGlnCysPheIlePh 218
Db 649 CCTAAGGCGCTGAAGAGGATGCCGCC-----TAGACTGCTTCACTT 693
Oy 218 eHisAspValaAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnPr 238
Db 694 CAGCGATGTGACCTGTGCCATGATGACCGCAACCTATACCGCTGCGGAGACCAACC 753
Oy 238 oArgHisMetSerValAlaIleAspIlePheAsnTyrIleLeuProTyrSerAlaIlePh 258
Db 754 CCGCCACTTTCCTGATGTCATGACCAAGTTTGCTTCCGGCTTCCCTATGCTGACTT 813
Oy 258 eGlyGlyIleSerAlaLeuThrIleAspHisIleuLeuIleAsnGlyPheSerAsnAs 278
Db 814 TGGAGGTGTGTAGAGCTGATGAGCTCATGTTCTGAGAAATCAATGCTTCCCAATGA 873
Oy 278 PheTTrpGlyTTrpGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLe 298
Db 874 GTAACGGGGCTGGGGTGGCGAGATGATGACATCTTCAACCGGATCTCCCTACGCGGAT 933
Oy 298 uIleValSerArgTyrProThrGlnIleAlaArgTyrIleMetIleIleHisSerThrGln 318
Db 934 GAAAGATCTCAACCGACATCCGATCCGCGCTAACCCGATGATCAACGACGCGCA 993
Oy 318 uAlaThrAsnProValAsnIleCysArgTyrIleIleMetGlyGlnThrIleArgArgTr 338
Db 994 CAAAGCATTAACGAACCTTAACGAGTTTACCAAGATTCAAAACAGAACTGACCAT 1053
Oy 338 pThrArgAspGlyLeuSerAsnLeuIleTyrIleLeuValaAsnLeuGluLeuIleAspProle 358
Db 1114 CTTCAACCAATATCAAGTGAACATT 1138
```

```
Db 1054 GAAGCGGACGCGATTGGGTCAGTGCGGTACCAAGGCTTGTGAGAGTGTCCGCAACCACT 1113
Oy 358 uTyrThrArgAlaValaAspLeu 366
Db 1114 CTTCAACCAATATCAAGTGAACATT 1138

Search completed: January 9, 2006, 06:39:01
Job time : 1018 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 9, 2006, 04:21:46 : Search time 982 Seconds
(without alignments)
284.350 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRIHVARUKSLVLCAV.....VLLKCKRRRLRDPYCF 383

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4637633 seqs, 364532575 residues
Total number of hits satisfying chosen parameters: 9275266

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10661430/runat_06012006_123139_6385/app_query.fasta.1.519
-DB=Published_Applications_NA_New -OFMT=fasta -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=500 -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0
-MAXLEN=200000000 -USFR=US10661430 @CGN 1.1.122 @runat_06012006_123139_6385
-NCPUG=6 -ICPU=3 -NO MMAP -LARGESUBJECT -NEG_SCORES=0 -WAIT -DSFBLOCK=100
-LOGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA_New.*

1: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	642.5	31.6	1964	US-11-102-240-41
2	635.5	31.3	951	US-11-136-527-1055
3	635.5	31.3	951	US-11-136-527-1055
4	583	28.7	4646	US-10-955-054A-120
5	331.5	16.3	2507	US-10-750-185-30398
6	331.5	16.3	2507	US-10-750-623-30398
7	313	15.4	16822	US-10-995-561-13325
8	313	15.4	18705	US-10-995-561-13401

9	163	8.0	1728	6	US-10-750-185-43601	Sequence 43601, A
10	163	8.0	1728	6	US-10-750-623-43601	Sequence 43601, A
11	154	7.6	201	6	US-10-995-561-22169	Sequence 22169, A
12	154	7.6	201	6	US-10-995-561-68468	Sequence 68468, A
13	136.5	6.7	3458	6	US-10-750-185-34973	Sequence 34973, A
14	136.5	6.7	3458	6	US-10-750-623-34973	Sequence 34973, A
15	109.5	5.4	3745	7	US-11-000-688-1051	Sequence 1051, Ap
16	100.5	4.9	3877	7	US-11-102-240-71	Sequence 71, Appl
17	87	4.3	90572	7	US-11-124-366A-2900	Sequence 2900, Ap
18	84	4.1	3148	6	US-10-793-626-4059	Sequence 4059, Ap
19	83	4.1	1521	7	US-11-134-241-30	Sequence 30, Appl
20	83	4.1	1812	7	US-11-134-241-42	Sequence 42, Appl
21	82.5	4.1	3896	6	US-11-134-241-44	Sequence 44, Appl
22	82.5	4.1	777	6	US-10-793-626-235	Sequence 235, Ap
23	82.5	4.1	1054	6	US-10-750-185-41517	Sequence 41517, A
24	82.5	4.1	1054	6	US-10-750-623-41517	Sequence 41517, A
25	82.5	4.1	3446	6	US-10-793-626-3943	Sequence 3943, Ap
26	82.5	4.1	3827	6	US-10-793-626-3598	Sequence 3598, Ap
27	82.5	4.1	3919	6	US-10-793-626-3814	Sequence 3814, Ap
28	82.5	4.1	57489	6	US-10-927-466-100	Sequence 100, App
29	82	4.0	1212	6	US-10-821-234-638	Sequence 638, App
30	82	4.0	1311	7	US-11-000-688-484	Sequence 484, App
31	82	4.0	4665	7	US-11-136-527-3309	Sequence 3309, Ap
32	81.5	4.0	3514	6	US-10-750-185-44589	Sequence 44589, A
33	81.5	4.0	3514	6	US-10-750-623-44589	Sequence 44589, A
34	80.5	4.0	2475	6	US-10-131-826A-195	Sequence 195, App
35	80.5	4.0	6243	6	US-10-821-234-788	Sequence 788, App
36	80	3.9	810	7	US-11-134-241-26	Sequence 26, Appl
37	80	3.9	1101	7	US-11-134-241-38	Sequence 38, Appl
38	80	3.9	3405	6	US-10-518-599-1	Sequence 1, Appl1
39	80	3.9	3732	6	US-10-131-826A-71	Sequence 71, Appl1
40	80	3.9	3900	7	US-11-197-721-42	Sequence 42, Appl1
41	79.5	3.9	1288	6	US-10-750-185-54909	Sequence 54909, A
42	79.5	3.9	1288	6	US-10-750-623-54909	Sequence 54909, A
43	79	3.9	2526	6	US-10-750-185-33449	Sequence 33449, A
44	79	3.9	2526	6	US-10-750-623-33449	Sequence 33449, A
45	79	3.9	190882	7	US-11-121-086-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-11-102-240-41
Sequence 41, Application US/11102240
Publication No. US20050260647A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey J.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESSION
FILE REFERENCE: ESOPHAGEAL TUMOR
CURRENT APPLICATION NUMBER: US/11/102,240
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: 10/063662
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 199-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 41
LENGTH: 1964
TYPE: DNA
ORGANISM: Homo Sapien
US-11-102-240-41
Alignment Scores: 5.76e-69 Length: 1964
Pred. No.:

Score: 642.50 Matches: 143
 Percent Similarity: 53.37% Conservative: 55
 Best Local Similarity: 38.54% Mismatches: 124
 Query Match: 31.63% Indels: 49
 DB: 7 Gaps: 8

US-10-661-430-1 (1-383) x US-11-102-240-41 (1-1964)

```

QY 5 HtLeuAlaValAla---ArgLeuYsSerLeuLeuValLeuCysAlaValLeuLeu--- 22
   |||||:::|||||
Db 140 CACCTTCCACAAATCCATTCCTGCTTGTGACTTGTGCTGACAGTGGTGGG 199
   |||||:::|||||
QY 23 -----LeuValHtAlaMetLeuYsLeuProSerLeuYsGlu 36
   |||||:::|||||
Db 200 TGGGCGACAGTAAGTACTCTGCGGGGCGC---ATTCAAGAGATTCTTAAGCAAGAG 256
   |||||:::|||||
QY 37 AsnLeuThrIleGlySerThrLeuIleAlaSerValAspAlaMetGluAlaValLeu 56
   |||||:::|||||
Db 257 TTCATGGCTAATTTCCATAGACCCCTCATTTTG-----GGGAAGGAAAACTCTG 307
   |||||:::|||||
QY 57 GluAsnThrAlaSerThrSerAspAspLeuLeuAspThrTrpAsnSerThrPheSerPro 76
   |||||:::|||||
Db 308 ACTAATGAACATCCACAGAAAGATGAAGCTTGACAC----- 346
   |||||:::|||||
QY 77 IleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAsp 96
   |||||:::|||||
Db 346 ----- 346
   |||||:::|||||
QY 97 AsnGlnThrLeuGlnPheCysAsnGlnThrProProHtIleuValGlyProIleArgVal 116
   |||||:::|||||
Db 347 -----TGTCTTCTGTGCTCTTCACTTCAGAGCGCCAGACAGACTC 388
   |||||:::|||||
QY 117 PheLeuAspGluProAspPheYsThrLeuGluYsIleYsTrpAspThrHis----- 134
   |||||:::|||||
Db 389 ATTTTC---AAACCAAGTCTC---ACTTGGAAAGAGTACAGCAGAAATCCCAAGTG 442
   |||||:::|||||
QY 135 AlaGlyGluHtIleGlyMetProYsAspCysValAlaArgHtIleArgValAlaIleVal 154
   |||||:::|||||
Db 443 TCAGAGGCGCGGATGCGCCTCAGAAATGAAGCTTACAGAGGCGCGCATCTCGTT 502
   |||||:::|||||
QY 155 ProYsTrpAspArgGluAlaHtIleuArgIleMetLeuHtIleuHtIleuHtIleu 174
   |||||:::|||||
Db 503 CCCACCGGAGACAGAGAAACACCTGATGTACCTGTGGAGACATCTGCATCCCTTCCTG 562
   |||||:::|||||
QY 175 AlaYsGlnIleuAspYsAlaIlePheIleValGluGlnValAlaAsnGlnThrPhe 194
   |||||:::|||||
Db 563 CAGAGCGACAGCTGGATTTGGCATCTACGTCTACACAGCGGTGAAGTAAAGTTT 622
   |||||:::|||||
QY 195 AsnArgGlyYsIleuMetAsnValGlyYsAspValAlaSerArgLeuYsProTrpGln 214
   |||||:::|||||
Db 623 AATCGAACCAACTCTTGAATGTGGCTATCTAGAGCCCTCAAGAAATTTGGGAC 682
   |||||:::|||||
QY 215 CysPheIlePheHtIleAspValAspLeuProGluAspAspArgAsnLeuYsTrpCys 234
   |||||:::|||||
Db 683 TGCCTTAATATTCACGATGTGAGCTGTACCCGACGATGACTTTAACCTTTAACAGTGT 742
   |||||:::|||||
QY 235 ProIleGlnProArgHtIleMetSerValAlaIleAspYsPheAsnYsYsLeuProYs 254
   |||||:::|||||
Db 743 GAGGAGCATCCCAAGACTCTGGTGGTGGCAGAGACGACACTGGGTACAGGTATCGTTAC 802
   |||||:::|||||
QY 255 SerAlaIlePheGlyYsIleSerAlaLeuThrYsAspHtIleuYsYsIleAsnGly 274
   |||||:::|||||
Db 803 AGTCGATATTTTGGGGGTGTACTGCGCCTAAGCAGAGAGCGATTTCAGGTGAATGGA 862
   |||||:::|||||
QY 275 PheSerAsnAspPheTrpGlyYsTrpGlyYsGluAspAspLeuAlaThrArgThrSer 294
   |||||:::|||||
Db 863 TTCCTTAACAACATGCTGGGAGTGGGAGCGCAGACATGACCTCAGACTCAGGTTGAG 922
   |||||:::|||||
QY 295 MetAlaGlyLeuYsValSerArgYsProThrGlnIleAlaArgYsYsMetIleYs 314
   |||||:::|||||
Db 923 CTCGAAGAATGAATTTCCCGGCGCCCTGCTTAAGTGGGTAAATTTACATAGTCTTC 982
   |||||:::|||||
QY 315 HisSerThrGluAlaThrAsnProValAsnYsCysArgYsYsIleMetGlyGlnThr 334
   |||||:::|||||

```

```

Db 983 CACACTAGACAAAGCGCATAGGGAACGAGAACGATGAAGCTCTTACACCAAGTG 1042
   |||||:::|||||
QY 335 LysArgArgTrpThrArgAspGlyLeuSerAsnLeuYsTrpYsLeuValAsnLeuGlu 354
   |||||:::|||||
Db 1043 TCACGAGTCTGGAAGAACATGGGTGAGTAGTGTCTTATTAATTAATATCTGGA 1102
   |||||:::|||||
QY 355 LeuYsProLeuYsTrpThrArgAlaValAsp 365
   |||||:::|||||
Db 1103 CACAAATCTTTATATATATCAACATCACAGTGAGT 1135
   |||||:::|||||

RESULT 2
US-11-136-527-1055
; Sequence 1055, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIORITY FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIORITY FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1055
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-11-136-527-1055

Alignment Scores:
Pred. No.: 1,366-68 Length: 951
Score: 635.50 Matches: 117
Percent Similarity: 62.83% Conservative: 52
Best Local Similarity: 43.49% Mismatches: 99
Query Match: 31.29% Indels: 1
DB: 7 Gaps: 1

US-10-661-430-1 (1-383) x US-11-136-527-1055 (1-951)
QY 99 ThrLeuGlnPheCysAsnGlnThrProProHtIleuValGlyProIleArgValPheLeu 118
   |||||:::|||||
Db 133 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 192
   |||||:::|||||
QY 119 AspGluPro---AspPheYsThrLeuGluYsIleYsTrpAspThrHisAlaGly 137
   |||||:::|||||
Db 193 AATATTCCTGTGATCTGAGCTTTTGGCAAGAAAGAACCAAGATGAAGTGGGGG 252
   |||||:::|||||
QY 138 HtIleYsMetProYsAspCysValAlaArgHtIleArgValAlaIleValProYsArg 157
   |||||:::|||||
Db 253 CGTTACTCTCCCAAGGCMCTATCTCCCTCCCAAGAGTGCATCTATCCCATTCCT 312
   |||||:::|||||
QY 158 AspArgGluAlaHtIleuArgIleMetLeuHtIleuHtIleuHtIleuHtIleu 177
   |||||:::|||||
Db 313 AACCGGACAGGACCTCAAAATAGCTGGTGTATTTGGCATGCAAGTCTTCAGCGCAG 372
   |||||:::|||||
QY 178 GluLeuAspYsAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGly 197
   |||||:::|||||
Db 373 CAATCGACTATGGCATCTACGTCATCATCAAGCTGGAGACCAATGTTTATCGAGCT 432
   |||||:::|||||
QY 198 LysLeuMetAsnValGlyYsAspValAlaSerArgLeuYsProTrpGlnCysPheIle 217
   |||||:::|||||
Db 433 AAGCTGCTCAAGCTTGGCTTTCAAGAGGCTTGAAGACATATGACATCAACTGCTTGG 492
   |||||:::|||||
QY 218 PheHtIleAspValAspLeuProGluAspAspArgAsnLeuYsTrpCysProIleGln 237
   |||||:::|||||
Db 493 TTCAGATATGAGCTCATTTCAATGATGACCATATATCCACAGAGTGCCTTTTCAGAG 552
   |||||:::|||||
QY 238 ProArgHtIleMetSerValAlaIleAspYsPheAsnYsYsLeuProYsProYsAlaIle 257
   |||||:::|||||
Db 553 CCAAGGCAATTTCTGTGCAATGACAAAGTTGGGTTTGAAGCTTGAAGTTCATAT 612
   |||||:::|||||

```

```

QY 258 PheGlyIleSerAlaLeuThrIysAspHisLeuLysLysIleAsnGlyPheSerAsn 277
DB 613 TTTGGAGGTCGTCCGCTTCAAGTAAACAAGTTCCTTACATCAATGATTTCTTAT 672
QY 278 AspPheTPrGlyTPrGlyGluAspAspLeuAlaThrArgThrSerMetAlaGly 297
DB 673 AATTCTGGGCTGGGAGGAGAGATGATGACATTTTAAACAGATTGATCAATAAGGC 732
QY 298 LeuLysValSerArgTyProThrGlnIleAlaArgTyLysMetIleLysHisSerThr 317
DB 733 ATGCTATATATCAAGCCCAATGCTGTGAGCGGTGCGATCATATCCGACATCAAGA 792
QY 318 GluAlaThrAspProValAsnLysCysArgTyLysIleMetGlyGlnThrIysArgArg 337
DB 793 GACAAAGTTTAATGAGCCCAACCTCAGAGGTTTGAACCGGATGCACATCAAAAGAAAC 852
QY 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyLysIleValAsnLeuGluLeuLysPro 357
DB 853 ATGCCGCTTGATGTTTGAACCTTACTTACAGGTTGTGACATACAGAGATACCCG 912
QY 358 LeuTyThrArgAlaValAlaAspLeu 366
DB 913 TTATATACCAAAATACAGTGCATC 939

RESULT 3
US-11-136-527-5151
; Sequence 5151, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5151
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-5151

Alignment Scores:
Pred. No.: 1,366-68 Length: 951
Score: 635.50 Matches: 117
Percent Similarity: 62.83% Conservative: 52
Best Local Similarity: 43.49% Mismatches: 99
Query Match: 31.29% Indels: 1
DB: 7 Gaps: 1

US-10-661-430-1 (1-383) x US-11-136-527-5151 (1-951)
QY 99 ThrLeuGlnPheCysAsnGlnThrProPheHisLeuValGlyProIleArgValPheLeu 118
DB 133 ACTCTGCTGCTTGGCTTGGAGAGTCCCGCTGCTGTTGGCCATGATGATGACTTT 192
QY 119 AspGluPro--AspPheLeuThrLeuGluLysIleTyProAspThrHisAlaGlyGly 137
DB 193 AATATATTCCTGCTGATCTGGAGCTTTGGCAAGAAAGAACCCAKAGTAAAGATGAGCTTT 252
QY 138 HisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIleValProTyArg 157
DB 253 CGTACTTCCCAAGGCTGATCTCCCTCACAAGGTGGCCATCATATCCATTCCT 312
QY 158 AspArgGluAlaHisIleLeuArgIleMetLeuHisAsnLeuHisSerIleLeuAlaLysGln 177
DB 313 AACCGGACGAGACCTCAAAATATCTGGCTGTATATTGCTTACCATCTTCAGGCGCCAG 372
QY 178 GlnLeuAspTyAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGly 197

```

```

DB 373 CAACCTGACCTATGCGATCTACGTATCATACAGGCTGGAGACACCATGTTTATGAGCT 432
QY 198 LysLeuMetAsnValGlyTyArgAspValAlaSerArgLeuTyProTPrGlnCysPheIle 217
DB 433 AAGCTGCTCAACGTTGGCTTCAAGAGGCTTGAAGAAGCTATGACTCAACAGCTTTTGG 492
QY 218 PheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThrCysProIleGln 237
DB 493 TTCAGTATGAGGACCTCATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 552
QY 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyLysLysProTySerAlaIle 257
DB 553 CCAAGGACATATTTCTGTCGCAATGACAGATTCGGGTTTAACTGCTTCAAGTTCATAT 612
QY 258 PheGlyIleSerAlaLeuThrIysAspHisLeuLysLysIleAsnGlyPheSerAsn 277
DB 613 TTTGGAGGTCGTCCGCTTCAAGTAAACAAGTTCCTTACATCAATGATTTCTTAT 672
QY 278 AspPheTPrGlyTPrGlyGluAspAspLeuAlaThrArgThrSerMetAlaGly 297
DB 673 AATTCTGGGCTGGGAGGAGAGATGATGACATTTTAAACAGATTGATCAATAAGGC 732
QY 298 LeuLysValSerArgTyProThrGlnIleAlaArgTyLysMetIleLysHisSerThr 317
DB 733 ATGCTATATATCAAGCCCAATGCTGTGAGCGGTGCGATCATATCCGACATCAAGA 792
QY 318 GluAlaThrAspProValAsnLysCysArgTyLysIleMetGlyGlnThrIysArgArg 337
DB 793 GACAAAGTTTAATGAGCCCAACCTCAGAGGTTTGAACCGGATGCACATCAAAAGAAAC 852
QY 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyLysIleValAsnLeuGluLeuLysPro 357
DB 853 ATGCCGCTTGATGTTTGAACCTTACTTACAGGTTGTGACATACAGAGATACCCG 912
QY 358 LeuTyThrArgAlaValAlaAspLeu 366
DB 913 TTATATACCAAAATACAGTGCATC 939

RESULT 4
US-10-955-054A-120
; Sequence 120, Application US/10955054A
; Publication No. US20050266420A1
; GENERAL INFORMATION:
; APPLICANT: PUSZTAI, LAJOS
; APPLICANT: SYMMANS, W. FRASER
; APPLICANT: HESS, KENNETH R.
; APPLICANT: AYERS, MARK
; APPLICANT: STEC, JAMES
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
; FILE REFERENCE: UYXC:880US
; CURRENT APPLICATION NUMBER: US/10/955,054A
; PRIOR FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 120
; LENGTH: 4646
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-955-054A-120

Alignment Scores:
Pred. No.: 5,976-61 Length: 4646
Score: 583.00 Matches: 127
Percent Similarity: 54.22% Conservative: 53
Best Local Similarity: 38.25% Mismatches: 136
Query Match: 28.71% Indels: 16
DB: 6 Gaps: 7

US-10-661-430-1 (1-383) x US-10-955-054A-120 (1-4646)
QY 39 ThrIleGlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsn 58

```



```

Db      287 ACAATCGTCTCAGCTTTATGAGACAGTCTTCGAGTCTTATGCCAAGAGAACAC 346
Qy      59 ThrAlaSerThrSerAspSerLeuAspThrTrpAsnSer---ThrPheSerProIle 77
Db      347 AGTGAATGACTCAATTTATCTCTTGACTGAAACACAGTAAACCTTC----- 397
Qy      78 SerGluValAsnGlnThrSerPheMetGluAspIleArgProIleuPheProAspAsn 97
Db      398 -----CTCGAAACTACAACTTTCTTCCGAA-----GACTTCACTACTTTGCAAC 445
Qy      98 GlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgValPhe 117
Db      446 CATACC-----TGCCCTGAAGAAGCTCTTCATGAAAGGCCCAATAGACAAAC 496
Qy      118 LeuAspGluProAspPheLeuSerThrLeuGluValLeuIleTyr-----ProAspThrHis 134
Db      497 ATGAGTGAATTGGAATTGATTTACATTCATGAATCTTCTCCAAAGAACCCACCATCAG 556
Qy      135 AAglyGlyHisGlyMetProlyAspCysValAlaArgHisArgValAlaIleIleVal 154
Db      557 CTGGAGAGTCACTGGAAGCTTCTGATTCATGATGCTCGGTGGAAGGTGGGATCTTATC 616
Qy      155 ProTyrArgAspArgGluAlaHisLeuValGlyLeuMetLeuHisAsnLeuHisSerLeu 174
Db      617 CCTTCCGGAACCCGACAGACGACCTCCAGTCTGTTCAAGACCTGCTTCCCATGCTC 676
Qy      175 AlaLeuGlnGlnLeuAspTyrAlaIlePheIleValGlnGlnValAlaAsnGlnThrPhe 194
Db      677 CAGGCGCAGGCTTCGATTCGATTTATGTGTTGAACAGTGGTATCCCAACCTT 736
Qy      195 AsnArgGlyLeuSerMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGln 214
Db      737 AATCGAGCCATGCTTTCAACGTTGCTTCAAGAGCATGAAGAAGCTTGATGGAC 796
Qy      215 CysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrTrpCys 234
Db      797 TGTGTTGTTTTCATGATGTGATGATCATACCCGAAAGTATGTCACATTTATGTGATGT 856
Qy      235 ProIleGlnProArgHisMetSerValAlaIleAspArgPheAsnTyrIleLeuProTyr 254
Db      857 GGAACAGTCCGAGGCAATTTGGCAACCAATTGGATATGATATGATCTTCTTCTTAT 916
Qy      255 SerAlaIlePheGlyGlyIleSerAlaLeuThrIleAspHisIleuValIleAsnGly 274
Db      917 ACGGAGTCTTTCGCGGAGTGAAGTGGCTTAACAGTGAACAATTCGGAATAATCAATGCG 976
Qy      275 PheSerAspAspPheTrpGlyTyrGlyGlyGluAspAspLeuAlaThrArgThrSer 294
Db      977 TTTCTTAATGCTTTCGCGGTTGGGGTGGAGAAGATGACGACCTCTGGAACAGAGTACAG 1036
Qy      295 MetAlaGlyLeuValSerArgTyrProThrGlnIleAlaArgTyrIleMetIleLys 314
Db      1037 AATGCAAGCTATTCTGTGAAGCGGCGCAGAGGCTGACACAGAAAGTACAGTCCATTCT 1096
Qy      315 HisSerThrGluAlaThrAspProValAsnLysCysArgTyrIleMetGlyGlnThr 334
Db      1097 CATCACATGAGAGAGAACTCAAGTTT---CTTGAAAGTATGCTCTGCTGAGAGAACTCA 1153
Qy      335 LysArgArgTrpThrArgAspGlyLeuSerAsnLeuLysTyrIleValLeuValAsnLeu 354
Db      1154 AAGAAACGGGCAAGGCTGATGGCTCAACAACCTGAACTAC---TTTGCAAAACATCA 1210
Qy      355 LeuLysProLeuTyrThrArgAlaValValAspLeu 366
Db      1211 TACAGCGCTTGTATAAAAACATTAACCTGCAACTG 1246

```

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30398
; LENGTH: 2507
; TYPE: DNA
; ORGANISM: Bovine 19866881075752
; US-10-750-185-30398

Alignment Scores:
Pred. No.: 5,31e-30 Length: 2507
Score: 331.50 Matches: 72
Percent Similarity: 50.53% Conservative: 24
Best Local Similarity: 37.89% Mismatches: 50
Query Match: 16.32% Indels: 44
Gaps: 3

US-10-661-430-1 (1-383) x US-10-750-185-30398 (1-2507)
Qy      142 LysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAla 161
Db      583 GAAGTGTCTTCTCTTCATCTCTTATCCCATACCTGCTTACTTCCCTCATTC 642
Qy      162 HisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyr 181
Db      643 ACTTTATGCTCTCTTTCAC-----GCCACTATAGACTTAATTGG 684
Qy      182 AlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsn 201
Db      685 TCCTCTCATATTATCTTCAGAGCTGGAATGGAACATTTAAGAGGCAAGCTCGTAAT 744
Qy      202 Val-GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVal 221
Db      745 GTTCGGGGTCCCGAGGCTCGCAGATGAGAAATGGAGCTGCTTCTTCTGATGATGT 804
Qy      221 AspLeuLeuProGluAspAspArgAsnLeuTyrTrpCys---ProIleGlnProArgHis 240
Db      805 GGAATCTCTGCGAGAGAACACCAACCTGTATGTGTGGATCCCGAGAGCTCGGCA 864
Qy      240 smetSerValAlaIleAspLysPheAsnTyrIleLys----- 251
Db      865 CGTTGCTGTTCGATACAAATTTGGATACAGTGAAGGCAACAGAGGCTACTGGAAG 924
Qy      251 ----- 251
Db      925 CAGGATATTCACCATCACTGGAAGAAGAGGAGTGGCTCCAAAAGCTGTGAATC 984
Qy      252 -----LeuProTyrSerAlaIlePheGlyGlyIleSerAlaIle 264
Db      985 CAAGATCTTTCTTCCCGAGCTCCGATGCCAGTACTTTGGGGGGTCTCAGGGCT 1044
Qy      264 uThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspPheTrpGlyTyrGlyGly 284
Db      1045 CACTCTGACCAATACCTGAAGATGATGGCTTCCCAATGATACCTGGGGGCTGGGGTGG 1104
Qy      284 YGluAspAspAspLeuAlaThrArgThr 293
Db      1105 TGAGGATGACGACATTTGCTACAGTCA 1132

```

RESULT 5
 US-10-750-185-30398
 ; Sequence 30398, Application US/10750185
 ; Publication No. US20050260603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.

RESULT 6
 US-10-750-623-30398
 ; Sequence 30398, Application US/10750623
 ; Publication No. US20050287531A1

GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30398
; LENGTH: 2507
; TYPE: DNA
; ORGANISM: Bovine 19866881075752
US-10-750-623-30398

Alignment Scores:
Pred. No.: 5,316-30 Length: 2507
Score: 331.50 Matches: 72
Percent Similarity: 50.53% Conservative: 24
Best Local Similarity: 37.89% Mismatches: 50
Query Match: 16.32% Indels: 44
DB: 6 Gaps: 3

US-10-661-430-1 (1-383) x US-10-750-623-30398 (1-2507)

QY 12 LysAspCysValAlaAlaHisArgValAlaIleIleValProTyrArgAspArgIleVal 161
DB 563 GAAATGTCCTTCTTCTTATCCCAATACCTGTTGACTTCCCTCCATTC 642
QY 162 HisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyr 181
DB 643 ACTTTATGCTTCTTCTCCAC-----GCCACCTATAGACTTGATTGG 684
QY 182 AlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgIlyLysLeuMetAsn 201
DB 685 TCCTCTCCATATATCTTCCAGGCTGGAATGGAACATTTAACAGGGCAAAAGCTGCTGAT 744
QY 202 Val-GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVal 221
DB 745 GTTCGGGGTCCGAGAGCCCTGCGAGATGAGAAATGGAGCTGTTGTTCTGCATGATGT 804
QY 221 AspleuLeuProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHis 240
DB 805 GGATCTCCGCGCAGAGAACACCACTGTATGTGTCCGATCCCGCAGACCTCGGCA 864
QY 240 sMetSerValAlaIleAspLysPheAsnTyrLys----- 251
DB 865 CGTTCCTGTCGACGAAACAAATTGGATACAGTAGAGGCGACAGAGGTAATCGAAG 924
QY 251 ----- 251
DB 925 CAGGGTATTCCACCATTCAGTGGAAAGAGGAGATGCTCCCAAAAGCTGTGGAATC 984
QY 252 -----LeuProTyrSerAlaIlePheGlyIleSerAlaIle 264
DB 985 CAAGATCTTCTTCTCCAGGCTCCCGATATCCCAAGTACTTTGGCGGGGTCTCAGCGGT 1044
QY 264 uThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTyrGlyIle 284
DB 1045 CACTCTCTGACAGTACCTGAAAGATGATGCTTCCCAATGAATACTGGGGCTGGGGTGG 1104
QY 284 YGluAspAspAspLeuAlaThrArgThr 293
DB 1105 TGAGGATGACGACATTGCTACCGAGTCA 1132
RESULT 7

US-10-995-561-13225/c
; Sequence 13225, Application US/10995561
; Publication No. US20050272054A1
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13225
; LENGTH: 16822
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(16822)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-;
US-10-995-561-13225

Alignment Scores:
Pred. No.: 2,216-26 Length: 16822
Score: 313.00 Matches: 68
Percent Similarity: 50.00% Conservative: 13
Best Local Similarity: 41.98% Mismatches: 36
Query Match: 15.41% Indels: 46
DB: 6 Gaps: 2

US-10-661-430-1 (1-383) x US-10-995-561-13225 (1-16822)

QY 185 IleValGlnGlnValAlaAsnGlnThrPheAsnArgIlyLysLeuMetAsnValGlyTyr 204
DB 13662 ATTATCTTCCAGGCTGGAATGGAACATTTAACAGGGCAAAAGCTGTAAGTTGGGGTG 13603
QY 205 AspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAspleuLeu 224
DB 13602 CGAGAGGCCCTGCGATGAGAGAGTGGAGCTGCTGTTCTTGCAAGTGTGACTCTTG 13543
QY 225 ProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHisMetSerVal 243
DB 13542 CCAGAAATAGACCAACATCTGTATGTGTGACCCCGGGAGCCCGGCATGTTGCCGT 13483
QY 244 AlaIleAspLysPheAsnTyrLys----- 251
DB 13482 GCTATGACCAAGTTTGTGATCAG-GTAGAGGCGATGGAAGGTACTGGGAAAACAGRAAG 13424
QY 251 ----- 251
DB 13423 CCTACTATCTGTGAGAGAGAGAGCCCAAGCGATTGTGGGATGTGTCCCGCAGA 13364
QY 252 -----LeuProTyrSerAlaIlePheGly 259
DB 13363 AGACTCGGGGTATGCAAGATCTTCTCTCCCTAGCTCCCGTACCCCAAGTACTTCGGA 13304
QY 260 GlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPhe 279
DB 13303 GAGCTCAGACATTAATCTGACCAAGTACCTGAAAGATGATGCTTCCCAATGAATAC 13244
QY 280 TrrpGlyTrrpGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLys 299
DB 13243 TGGGGCTGGGGTGGTGGAGATGACACATTGCTACCAAGTACGAGTCTTCCGCACTTCT 13184
QY 300 ValSer 301
DB 13183 CCCTCC 13178
RESULT 8
US-10-995-561-13401
; Sequence 13401, Application US/10995561
; Publication No. US20050272054A1

Alignment Scores:	
Pred. No.:	1.35e-09
Score:	154.00
Percent Similarity:	64.00%
Best Local Similarity:	58.00%
Length:	20
Matches:	29
Conservative:	3
Mismatches:	18

FILE REFERENCE: MM1100-2

```
/ CURRENT APPLICATION NUMBER: US/10/750,185
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 34973
/ LENGTH: 3458
/ TYPE: DNA
/ ORGANISM: Bovine 19866881077545
US-10-750-185-34973

Alignment Scores:
Pred. No.: 1.87e-05 Length: 3458
Score: 136.50 Matches: 30
Percent Similarity: 64.91% Conservative: 7
Best Local Similarity: 52.63% Mismatches: 19
Query Match: 6.72% Indels: 2
DB: Gaps: 1

US-10-661-430-1 (1-383) x US-10-750-185-34973 (1-3458)

QY 188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyr---AspVal 206
Db 3288 CAGCACGGTGAAGACACACCTTCAACCGGCCCAAGCTGCTCAACGGGTTTCTTAGAGGCA 3347

QY 207 AlaSerArgLeuTyrProTropGlnCysPheHisAspValAspLeuLeuProGlu 226
Db 3348 CTCAGAGAGACTCCACCTTCAACTGCTTCACTTCTAGTACGTGACCTGAGTCCCATG 3407

QY 227 AspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerVal 243
Db 3408 GATGACCGCAACTGTACCGCTGTGG-TACCAAGCCCGCCACCTTGGCATT 3457

RESULT 14
US-10-750-623-34973
/ Sequence 34973, Application US/10750623
/ Publication No. US20050287531A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: PANTIN, Dennis
/ TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: MM1100-1
/ CURRENT APPLICATION NUMBER: US/10/750,623
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 34973
/ LENGTH: 3458
/ TYPE: DNA
/ ORGANISM: Bovine 19866881077545
US-10-750-623-34973

Alignment Scores:
Pred. No.: 1.87e-05 Length: 3458
Score: 136.50 Matches: 30
Percent Similarity: 64.91% Conservative: 7
Best Local Similarity: 52.63% Mismatches: 19
Query Match: 6.72% Indels: 2
DB: Gaps: 1

US-10-661-430-1 (1-383) x US-10-750-623-34973 (1-3458)

QY 188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyr---AspVal 206
Db 3288 CAGCACGGTGAAGACACACCTTCAACCGGCCCAAGCTGCTCAACGGGTTTCTTAGAGGCA 3347
```

```
QY 207 AlaSerArgLeuTyrProTropGlnCysPheHisAspValAspLeuLeuProGlu 226
Db 3348 CTCAGAGAGACTCCACCTTCAACTGCTTCACTTCTAGTACGTGACCTGAGTCCCATG 3407

QY 227 AspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerVal 243
Db 3408 GATGACCGCAACTGTACCGCTGTGG-TACCAAGCCCGCCACCTTGGCATT 3457

RESULT 15
US-11-000-688-1051
/ Sequence 1051, Application US/11000688
/ Publication No. US20050287544A1
/ GENERAL INFORMATION:
/ APPLICANT: BERTUCCI, Francois
/ APPLICANT: BERTUCCI, Francois
/ APPLICANT: HOUIGATTE, Remi
/ APPLICANT: BIRNBAUM, Daniel
/ TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
/ FILE REFERENCE: 1423-R-03
/ CURRENT APPLICATION NUMBER: US/11/000,688
/ CURRENT FILING DATE: 2004-12-01
/ PRIOR APPLICATION NUMBER: US 60/525,987
/ PRIOR FILING DATE: 2003-12-01
/ NUMBER OF SEQ ID NOS: 1596
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1051
/ LENGTH: 3745
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial sequences:iprimer
/ NAME/KEY: misc feature
/ LOCATION: (1)-(3745)
/ OTHER INFORMATION: chondroitin sulfate galnact-2(GALNACT-2)
US-11-000-688-1051

Alignment Scores:
Pred. No.: 0.0497 Length: 3745
Score: 109.50 Matches: 98
Percent Similarity: 34.57% Conservative: 61
Best Local Similarity: 21.30% Mismatches: 163
Query Match: 5.39% Indels: 138
DB: Gaps: 20

US-10-661-430-1 (1-383) x US-11-000-688-1051 (1-3745)

QY 30 LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIleAlaAspVal 49
Db 747 AAACATCCAGATGATGAGTATGGGTCATTCCCTTGAAGATTTCCTTA-----ATG 797

QY 50 AspAlaMetGluAlaValLeuGlyAsnThr-----Ala 60
Db 798 AAAGTATTTCAATTGAAATGGATGCTCCTCCATCCGATCTGAAGAAAGCCAGTTGAAAA 857

QY 61 SerThrSerAspAspLeuAspThrTyrPasnSerThrPheSerProIleSerGluVal 80
Db 858 GACAAACGAGATGAATTTGGTGAAGATTATGAAGCGGCTGGAGGTCATTAATATTCCT 917

QY 81 AsnGlnThrSerPheMetGluAspIle-----ArgProIleLeuPhePro 95
Db 918 GATGACGATGATGAACAGAAAGATGAGAGGATCCCTTGGAGAGAAAGATATTTAAT 977

QY 96 AspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArg 115
Db 978 GAAATGACTTTCGTAGAAGGTTATATTCGACTGAAGAGATGAAGGACACAGTATGAA 1037

QY 116 ValPheLeuAspGluProAsp-----Phe 123
Db 1038 CTCCTTTTAAAGAAAGACACCTTAGAGATATAGACATGTGACCCCTTCCGCGCTTTT 1097

QY 124 LysThrLeuGluLysIleTyrProAspThrHisAlaGlyGlyHisValMetProLysAsp 143
```

```
DB 1098 GGACCTTCATGAAAGTGAAGATGATTGAC----- 1133
QY 144 CysValAlaArgHisArgValAlaIleIleValProTyrArgAspArg---GluAlaHis 162
DB 1134 ---ATCATAGATCAATTATTAATATCAATTGTGCCACTTGTGAAGAAGTGAAGCATTTT 1190
QY 163 LeuArgIleMet-----LeuHisAsn 169
DB 1191 GTACAAATTTATGCAAGACTTCAGGAGTGTGTGATTCATCAAGACAAAGAGATTCTCTC 1250
QY 170 LeuHisSerLeuLeuAlaLysGlnLeuAspTyrAlaIlePheIleValGluGlnVal 189
DB 1251 ACAGTGTGTATTTTGTGAAGAAGACGTGTAAAGTCAAGTCTATCCTAGAAATCTGTCTC 1310
QY 190 Ala-----AsnGlnThrPheAsnArg 196
DB 1311 ACCAGTAGCTCAATTTCACAATTACACCTTGCTCATTTGATGATGAAAGAAATTTAATCGT 1370
QY 197 GlyLysLeuMetAsnValGly-----TyrAspValAlaSerArgLeuTyrProTyr 213
DB 1371 GGAACGAGACTAAATATGTGGTGGCCGAGCTTGGAGACAGAGAGGCTTTG----- 1421
QY 214 GlnCysPheIlePheHisAspValAsp----- 222
DB 1422 ---ATGTTTTTCTGTGATGTGATATCTAATTCTCAGCCGAAATTCCTTAACAGCTGC 1475
QY 223 ---LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMet 241
DB 1476 CGGTTAAATGCTGAGCAGAGTAAGAGGTGTTT---TACCCTGTGGTGTTCAGTCTTTAC 1532
QY 242 SerValAlaIleAspLysPheAsnTyrLysLeuPro----- 253
DB 1533 AATCTGCGCAATGTTATATGCCAACCAGAAAGTCCACCACTGTGAGACAGACGCTGTT 1592
QY 254 ---TyrSerAlaIlePheGly---GlyLysSerAlaLeuThrLys 266
DB 1593 CACAAAAGAGATTCTGCTTTTGGCGAGATTTTGGCTTTGGAATGACTTGTCAAGTATCGT 1652
QY 267 AspHisIleLeuLysValIleAsnGlyPheSerAsnAspPheTrpGlyTrpGlyGluAsp 286
DB 1653 TCAGATTTCCTGACCACTGTGGATTGACATGAGAGTGAAGAAGTTGGGGTGGAGAAAGAT 1712
QY 287 ---AspAspLeuAlaThrArgThrSerMetAla 296
DB 1713 GTTCATCTTTATGAAATACTTACATGCTGACCTGATTGTGATTCGACTCCGGTTCTT 1772
QY 297 GlyLeuLysValSerArgTyrProThrGlnIleAla-----ArgTyr 310
DB 1773 GGTCTTTTCCACCTCGCGCATGAAGAAGCGCTGTGATGAGCTGACCCCGAGCAGTAC 1832
QY 311 LysMetIle-----LysHisSerThrGluAlaThrAsn-ProValAsnLysCys-- 326
DB 1833 CCGATGTGCATCAGCTTAAGCATGAATGAGGCTCTCACCTCCACCTCGGGAATGCTG 1892
QY 327 ---ArgTyrLysIleMetGlyGlnThr 334
DB 1893 GTCTTCAGGAGAGAAATAGACAGCATCTTCATTAACAGGCAAT-----ACAGACAAAC 1946
QY 334 rLysArgArgTrpThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn----- 352
DB 1947 AGTGAAGCTGTGGTGAATCATTAATATGCGTTACTGTATGAAACCAAAAACAGCAC 2006
QY 353 -LeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeuLeuGlnLysAspCys 371
DB 2007 TATTTATTTAGCTTAATTCTACTTCCAGATGCAAGTGCCTTTTGGAGAGACATGT 2064
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:23:20 ; Search time 28 Seconds
(without alignments)
116.088 Million cell updates/sec

Title: US-10-661-430-1

Perfect score: 2031

Sequence: 1 MAFRLAVARLKSILVLCANV.....VDLEKDCRRRLRDPPTCF 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 61072 seqs, 8486849 residues

Total number of hits satisfying chosen parameters: 61072

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642.5	31.6	344	US-11-102-240-42	Sequence 42, Appl
2	100.5	4.9	532	US-11-102-240-72	Sequence 72, Appl
3	87	4.3	885	US-10-055-877-240	Sequence 240, Appl
4	87	4.3	885	US-10-055-877-241	Sequence 241, Appl
5	83	4.1	506	US-11-134-241-31	Sequence 31, Appl
6	83	4.1	603	US-11-134-241-43	Sequence 43, Appl
7	82	4.0	403	US-10-821-234-1490	Sequence 1490, Ap
8	80.5	4.0	552	US-10-131-826A-196	Sequence 196, Appl
9	80.5	4.0	2080	US-10-821-234-1600	Sequence 1600, Ap
10	80	3.9	269	US-11-134-241-27	Sequence 27, Appl
11	80	3.9	366	US-11-134-241-39	Sequence 39, Appl
12	78	3.8	258	US-10-793-626-236	Sequence 236, Appl
13	77.5	3.8	704	US-11-172-145-25	Sequence 25, Appl
14	77.5	3.8	965	US-11-172-145-4	Sequence 4, Appl
15	77.5	3.8	972	US-11-124-215-3	Sequence 3, Appl
16	77	3.8	458	US-11-016-503-16	Sequence 16, Appl
17	77	3.8	458	US-11-089-803-6	Sequence 6, Appl
18	77	3.8	458	US-11-149-738-2	Sequence 2, Appl
19	77	3.8	458	US-11-155-266-2	Sequence 2, Appl
20	77	3.8	2204	US-11-052-554A-174	Sequence 174, Appl
21	77	3.8	2335	US-10-821-234-1610	Sequence 1610, Ap
22	76.5	3.8	385	US-10-392-234A-28	Sequence 28, Appl
23	75	3.7	858	US-10-645-441-18	Sequence 18, Appl
24	75	3.7	858	US-10-645-441-20	Sequence 20, Appl
25	75	3.7	858	US-10-645-441-23	Sequence 23, Appl

26	75	3.7	896	US-10-467-657-7004	Sequence 7004, Ap
27	74.5	3.7	458	US-11-016-503-12	Sequence 12, Appl
28	74.5	3.7	458	US-11-089-803-2	Sequence 2, Appl
29	74.5	3.7	513	US-10-878-556A-112	Sequence 112, Appl
30	74	3.6	561	US-10-454-437-330	Sequence 330, Appl
31	72.5	3.6	430	US-11-016-503-17	Sequence 17, Appl
32	72.5	3.6	621	US-10-858-730-5	Sequence 5, Appl
33	72.5	3.6	621	US-10-632-150-28	Sequence 28, Appl
34	72.5	3.6	621	US-11-073-457-28	Sequence 28, Appl
35	72.5	3.6	621	US-11-074-460-28	Sequence 28, Appl
36	72.5	3.6	4868	US-11-044-111-24	Sequence 24, Appl
37	72	3.5	535	US-11-134-241-29	Sequence 29, Appl
38	72	3.5	631	US-11-134-241-41	Sequence 41, Appl
39	72	3.5	914	US-10-312-954-2	Sequence 2, Appl
40	72	3.5	1255	US-11-052-554A-273	Sequence 273, Appl
41	71.5	3.5	774	US-11-070-627-7	Sequence 7, Appl
42	71.5	3.5	1062	US-10-821-234-1079	Sequence 1079, Ap
43	71	3.5	376	US-10-454-437-412	Sequence 412, Appl
44	71	3.5	775	US-10-131-826A-120	Sequence 120, Appl
45	70.5	3.5	415	US-11-110-851-64	Sequence 64, Appl

ALIGNMENTS

```
RESULT 1
US-11-102-240-42
; Sequence 42, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William L.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 42
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-42
Query Match          31.6%  Score 642.5; DB 7; Length 344;
Best Local Similarity 38.5%  Pred. No. 3.9e-55;
Matches 143; Conservative 55; Mismatches 124; Indels 49; Gaps 8;

QY      5 HLAFA-RLKSLVLCVLL-----LVHAMIYKIPISYENLTGSSSTLIADVAMEAVL 56
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      8 HLSYFRLLLTLLTCLTVGVGATSNVYFGA-IQELPKAKFANPHKTLIL---GKGKTL 63
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      57 GNTASTSDLLDTMNSTSPISSEVNGTSFEMEDIRLPLPDNQTLOPCNQTPHLVGPPIRV 116
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      64 TNEASTKKEVELDN-----CPSVSPLYLGQSKTL 90
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      117 FLDEDPFTLEKIPYDTH--AGHGMPKDCVARRVATIIYVPRDEARLRLTLHLHSIL 174
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      91 IF-KPDL-TLEEVQANRVSRRGRPOECKALQVALIVPRNRKHLMTLEHLHPL 148
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      175 AKQOLDVAIFIVEQVANQTFNRGKLMMVGYDVASRLYPMQCFIFADVLLPEDENLYTC 234
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 149 ORQOLDYGIYVIHQAEKGKFNRAKLINVGYLEALKEENNDCFIHDVDLVPENDFNLYXC 208
;
Qy 235 PLOPHMSVAIDKFNKLYSAIFGGISALTKDHKKINGFSDNFGWGGEDDLATRRS 294
;
Db 209 EEPKHLVVGRRSTGRLRYSYGGVTLASREQFKNVGSNNVWGEGEDDLRLRYE 268
;
Qy 295 MAGLKVSRYPTQIARXKMTIKHSTEAATPNVKKCRKYMGTQKRRMTRDGLSNLYKLVNLE 354
;
Db 269 LORMKISRLLPVRGKTYMFIHTRDKGNENVAERMKLLHGVSRVTRDGLSSCSYKLVSYE 328
;
Qy 355 LKPLVTRAVD 365
;
Db 329 HNPLYINITVD 339
;
RESULT 2
US-11-102-240-72
; Sequence 72, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESSION
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 72
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-72
;
Query Match 4.9%; Score 100.5; DB 7; Length 532;
Best Local Similarity 22.1%; Pred. No. 0.043;
Matches 55; Conservative 34; Mismatches 93; Indels 67; Gaps 11;
;
Qy 145 VARRVAIVPYDREAHRLIMLNLHSL-----LAKQOLDYAIPIVEOVA 190
;
Db 258 MANTLINIVPLAKVDKFPQFQONPREMCIEDGRVHLTVYFGKEINVEVGILENTS 317
;
Qy 191 -----NQTFNRKLMNVGYDVASRLYPWQ---CFIHDVDLLPEDDRNLTY 233
;
Db 318 KANFNRFETIQNGEFSRCK---GLDVGARF--WGSNVLLFFCVDVITFISE--PLNT 370
;
Qy 234 C-----PIQRRHNSVALDKFNKLYPY---SAI-----FG-CISALTYK 266
;
Db 371 CRNLTPGKKVYFVLFSSQYNPGIYGHNDVAPLEQOLVIKKEGTGWRDFGCMTCQYR 430
;
Qy 267 DHLKKINGFSDNFGWGGEDDLATRTSMAGLKVSRYPTQIARXKMTIKHSTEAATPNVKK 326
;
Db 431 SDFINIGFPLDIKMGGEDEVHLRYKYLHSLNLYVRFPVAGLPH--LHMEKRCMDLTPRE 468
;
Qy 327 RYKIMGQTK 335
;
Db 489 QYKCMQSK 497
;
RESULT 3
US-10-055-877-240
; Sequence 240, Application US/10055877
```

```
; Publication No. US2005028241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Rattelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zehnusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patuturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eissen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 240
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-240
;
Query Match 4.3%; Score 87; DB 6; Length 885;
Best Local Similarity 18.9%; Pred. No. 1.8;
Matches 61; Conservative 45; Mismatches 82; Indels 134; Gaps 15;
;
Qy 114 IRVLEDPDEPKTEKTYRPTNAGHGMPKDCVARRHVAIVPYDREAHRLIMLNLHSL 173
;
Db 69 VRIFLTGRLPGLGKLYQF-----PRE-----LPBER-----AQVTR 103
;
Qy 174 LAKQOLDYAIPIVEOVANQTFNRKLMNVGYDVASRLYPWQCFIHDVDLLPEDDRNLTY 233
```

```
Db 104 LKQ-----AEEVNT-----WEKLNHSADWQKIDETLE- 134
Qy 234 CPIOGRHNSVAIDKKNYLPYSAIFG-----ISALTQDHUKKINGSFNDWGMG 284
Db 135 ---RLQELQEAIDELDLKRLQAEVKGSMQPVGDLLISL-ODHLEKVALR-----G 183
Qy 285 E-----DDDLATRTSMAGLKVSRYPTQI-----ARYKMIGHSTE----- 318
Db 184 EIAPIKENVSHVNDLARQLTTIGIQLSPYNLSTEDLNTFWLQVAVEDRQLHEAR 243
Qy 319 -----ATNPVNCCKRYKMGOTKAR-WTRDGLSNLKYLVNL----- 353
Db 244 DFGPASQHFLLSTSVQGPWEARISP-NKVPPYINHETQTCMDHPKMTLEYOSLADLNNVR 302
Qy 354 -----ELKPLYTRAVDDL 367
Db 303 FSAVRTAMKLRRLQKALCLDIL 324
```

RESULT 4

```
US-10-055-877-241
; Sequence 241, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchervnev, Vellizar
; APPLICANT: Zhong, Mel
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spyrek, Kimberly
; APPLICANT: Rateili, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zernusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patuturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eissen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Bolog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
```

```
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 241
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-241
```

Query Match 4.3%; Score 87; DB 6; Length 885;
Beet Local Similarity 18.9%; Pred. No. 1.8; Indels 134; Gaps 15;
Matches 61; Conservative 45; Mismatches 82

```
Qy 114 IRVFLDEPDPFTLEKIYPDTAGHGMPKDCVARRHVAIIVPYRDREAHRLIMLNHSL 173
Db 69 VRIFLTEQPLBGLKELYE-----PRE-----LPPEER-----AQNVTRL 103
Qy 174 LAKQQLDYAIFIVEQVANQTFNRGKLMNVGYDVASRLYPWOCFTFHDVDLPEDDRLTYT 233
Db 104 LKQ-----AEEVNT-----WEKLNHSADWQKIDETLE- 134
Qy 234 CPIOGRHNSVAIDKKNYLPYSAIFG-----ISALTQDHUKKINGSFNDWGMG 284
Db 135 ---RLQELQEAIDELDLKRLQAEVKGSMQPVGDLLISL-ODHLEKVALR-----G 183
Qy 285 E-----DDDLATRTSMAGLKVSRYPTQI-----ARYKMIGHSTE----- 318
Db 184 EIAPIKENVSHVNDLARQLTTIGIQLSPYNLSTEDLNTFWLQVAVEDRQLHEAR 243
Qy 319 -----ATNPVNCCKRYKMGOTKAR-WTRDGLSNLKYLVNL----- 353
Db 244 DFGPASQHFLLSTSVQGPWEARISP-NKVPPYINHETQTCMDHPKMTLEYOSLADLNNVR 302
Qy 354 -----ELKPLYTRAVDDL 367
Db 303 FSAVRTAMKLRRLQKALCLDIL 324
```

RESULT 5

```
US-11-134-241-31
; Sequence 31, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kleke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/11/134,241
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
```

PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 506
TYPE: PRT
ORGANISM: homo sapiens
US-11-134-241-31

Query Match 4.1%; Score 83; DB 7; Length 506;
Best Local Similarity 17.5%; Pred. No. 2;
Matches 49; Conservative 41; Mismatches 80; Indels 110; Gaps 11;

QY 72 STFSPISEVNTQTSFMDIRPIILFPDNTQLOFCNQTPPHLVGPIRVFLDEPDEFTLEK--- 128
DB 51 STIIPFHNEGWSLLRTVHSV-----NRSPELVAEIVLVDPSDREHLKKPLR 100
QY 129 ----IYP-----DTN--AGGHGMPK--DCVAH 148
DB 101 DYMALEPFSVRIILTKKREGILRTRMGLASVATGDIVITFLDSHCEANVMWLPPLLDRIAN 160
QY 149 RVAIVPYRDREAHRLIMLHNLHSLAKQQLDYAIFIVEQVANQTFNRGKLMNVGYDVA 208
DB 161 RKTIVCPMID-----VIDHDPFRYETQA-----GDAMRGAFD--- 192
QY 209 RLYPMQCFIFHDVLLPE---DDRNLVYTCPIQPRHMSVAIDKFNYKLPYSAIFGIGIAL 264
DB 193 ---WEMY-YKRIPIPELQKADPSDPFESPV-----MAGGLFAV 227
QY 265 TKDHLKKNFNSDFMGCGEDDLATRTSMAGLKVSRY 304
DB 228 DRKFWELGIGYDPLGLEIHWGSEQEISFKVMCGGRMEDIP 267

RESULT 6
US-11-134-241-43
Sequence 43; Application US/11134241
Publication No. US20050287566A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Hilbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilganowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kieke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT APPLICATION NUMBER: US/11/134,241
CURRENT FILING DATE: 2005-05-20
PRIOR APPLICATION NUMBER: US/10/364,774
PRIOR FILING DATE: 2003-02-11
PRIOR APPLICATION NUMBER: US/09/795,926
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 603
TYPE: PRT
ORGANISM: homo sapiens

US-11-134-241-43

Query Match 4.1%; Score 83; DB 7; Length 603;
Best Local Similarity 17.5%; Pred. No. 2.6;
Matches 49; Conservative 41; Mismatches 80; Indels 110; Gaps 11;

QY 72 STFSPISEVNTQTSFMDIRPIILFPDNTQLOFCNQTPPHLVGPIRVFLDEPDEFTLEK--- 128
DB 148 STIIPFHNEGWSLLRTVHSV-----NRSPELVAEIVLVDPSDREHLKKPLR 197
QY 129 ----IYP-----DTN--AGGHGMPK--DCVAH 148
DB 198 DYMALEPFSVRIILTKKREGILRTRMGLASVATGDIVITFLDSHCEANVMWLPPLLDRIAN 257
QY 149 RVAIVPYRDREAHRLIMLHNLHSLAKQQLDYAIFIVEQVANQTFNRGKLMNVGYDVA 208
DB 258 RKTIVCPMID-----VIDHDPFRYETQA-----GDAMRGAFD--- 289
QY 209 RLYPMQCFIFHDVLLPE---DDRNLVYTCPIQPRHMSVAIDKFNYKLPYSAIFGIGIAL 264
DB 290 ---WEMY-YKRIPIPELQKADPSDPFESPV-----MAGGLFAV 324
QY 265 TKDHLKKNFNSDFMGCGEDDLATRTSMAGLKVSRY 304
DB 325 DRKFWELGIGYDPLGLEIHWGSEQEISFKVMCGGRMEDIP 364

RESULT 7
US-10-821-234-1490
Sequence 1490; Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PC SEQ_genes Version 1.0
SEQ ID NO 1490
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1490

Query Match 4.0%; Score 82; DB 6; Length 403;
Best Local Similarity 21.5%; Pred. No. 1.8; Indels 116; Gaps 19;
Matches 76; Conservative 45; Mismatches 117;

QY 76 PISEVNTQTSFMDIRPIILFPDNTQLOFCNQTPPHLVGPIRVFLDEP---DEFTL----- 126
DB 63 PGSKVKKKEVEAVTIV-----ETPRMVVVGIVGYETRGILRTFKTFAEHI 110
QY 127 ----EKIYPTTAGS-----HGPKKC-----VABHVAII 153
DB 111 SDECKRRFYKNMHSKKKAKFTYCKKQODEGKKOLEKFSSMKKCYQVIRVIAHQMRL 170
QY 154 VPYRDREAHRLIMLHNLHSLAKQQLDYAIFIVEQ--VANQTFNRGKLMNV-----GY 204
DB 171 LPLRQKKAHL--MEIQVNGTVAEKIDMARERLEQVNPVQVGGODEMIDVIGTGTGKG 228
QY 205 -DVASRLYPWQCFIFHDVLLPEDDRNL--YTC-----PIQPR-----HMSVAI 245
DB 229 KGYTSR-----WTKKLPRTKTRGLRKVACIGAMHFAVAFSVARAGQKGYHRTET 280
QY 246 DKFNYYKLPYSAIFGIGIALTKDHLKKNFNSDFMGCGEDDLATRT--SMAGLKVSRY 303
DB 281 NKRIYKI-----GGGYLIKDGKLIKNNASTDY-----DLSDKSINPLGSP--VHY 323

QY 304 PROIARYKIKHSTATNP-VNCKRYKIMGOTRBRTRDGLSLTKYLVNLEIK 356
DB 324 GEVTDFWMLKGCVCVGTKKRVLTLRKSLVOTRR---ALEKIDUKFIDTTSK 373

RESULT 8

US-10-131-826A-196
; Sequence 196, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flivaotoff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131, 826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 196
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-196

Query Match 4.0%; Score 80.5; DB 6; Length 552;
Best Local Similarity 22.0%; Pred. No. 4;

Matches 54; Conservative 36; Mismatches 77; Indels 79; Gaps 14;

QY 82 QTSFMEDIRPILFPDNOQLQFCNQTPPHLVGPPIV---FLDEF-DPTELEKTPDTHAGG 137
DB 124 RSTLARTISVL-----NRTPHLRREILVDDSDNDPDDCKQLIK----- 164
QY 138 HGMFK-DCVARRHVAIIVPYRD-----EAHRLIMLHNLHSLAKQQLDYAI 183
DB 165 --LPKVKCLRNRRGQIVRSIRIGADIAGQTLTFLDSHCEVNRDWLQPLHLRVKEDYTR 222

QY 184 F---IVEQVANQTFN-----RGKLMNVGYDVASRLYPWQCFIFHDVLLPEDD-RNL 231
DB 223 YVCEPIDIDINDTFTYISASLRG-----GFD-WSLRFQW-----QSPFQKARL 269
QY 232 YTCPIQPRHMSVAIDKFNKYLPYSAIFGIGISALTYDHLKKTNGFENDFWGCGEDDDLAT 291
DB 270 --DPTEPIRTPI-----IAGGLFVIDKAMFIDYLGKYMDDMDIMWGENFEISF 314

QY 292 RTSNAG 297
DB 315 RVMWCG 320

RESULT 9

US-10-821-234-1640
; Sequence 1640, Application US/10821234
; Publication No. US2005025114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andermani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821, 234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462, 047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc seq_genes Version 1.0
; SEQ ID NO 1640
; LENGTH: 2080
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1640

Query Match 4.0%; Score 80.5; DB 6; Length 2080;
Best Local Similarity 21.7%; Pred. No. 26;
Matches 53; Conservative 31; Mismatches 73; Indels 87; Gaps 12;

QY 38 LTIGSSTLIADVDAMEAVLGNTASTSDLLDTWNSTPSPISVNOT-SFWE-DIRPILFP 95
DB 1450 LSPGEDVIL-DIDDEPIL---PIQEEFIDWMSKXFASIGREKGSYLEKDFDLKYV 1505
QY 96 DNG-----TLQFCN-----QTPPHLVGIRFLDEPDPKTEKTYPDTH 134
DB 1506 DTQLNVEAFEGLSDFCNTFKLYRGKTOTETEDPSVIG-----EFKGLFKIYP--- 1553
QY 135 AGGHGMPKDCVARRHVAIIVPYRD-----EAHRLIMLHNLHSLAK----- 176
DB 1554 -----LPED-----PALPMPRQFHQLAAGQFQECVRIIYVRAFGLOPKDPNGKCDPY 1602
QY 177 -----QQLDYAIFIVEQVANQTFNRGKLMNVGYDVASRLYPWQCFIFHDVLLP 225
DB 1603 IKISIGKKSVDQDNVYIPCTLEPVPFGKMFELCTPLPLEKDLKITLY-----DYDLIS 1654
QY 226 EDDR 229
DB 1655 KDEK 1658

RESULT 10

US-11-134-241-27
; Sequence 27, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.

```
APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kleke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT APPLICATION NUMBER: US/11/134,241
CURRENT FILING DATE: 2005-05-20
PRIOR APPLICATION NUMBER: US/10/364,774
PRIOR FILING DATE: 2003-02-11
PRIOR APPLICATION NUMBER: US/09/795,926
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 269
TYPE: PRT
ORGANISM: homo sapiens
US-11-134-241-27
```

```
Query Match 3.9%; Score 80; DB 7; Length 269;
Best Local Similarity 17.9%; Pred. No. 1.6;
Matches 50; Conservative 41; Mismatches 75; Indels 114; Gaps 12;
```

```
QY 72 STFSPISEVNTQSFMEDIRPLFPDNTQLQFCNQTPHVLGPIRVFLDEDPFTLEK--- 128
DB 51 STIIFPHNEGWSLRTVHSV-----NRSPPELVAEIVLVDDFSDEHLKPLE 100
QY 129 ----IYP-----DTH--AGGHGMPK--DCVARR 148
DB 101 DYVALPFSVRILRTKKREGILRTRMGASVATGDTVITFLDSHCANVNMPLPILDRIRRN 160
QY 149 RVAIIVPYRDREAHRLIMLNHLSLAKQOLDVAIFVEQVANQTFNRGKLMNVGYDVAS 208
DB 161 RKTIVCPMD-----VIDHDDFRYETQA-----GDAMRGAPD--- 192
QY 209 RLVPMOCFTFHDVLLPE---DDRNLYTCPIQPRHMSVAIDKENYKLPYSALFGISAL 264
DB 193 ----WEMV-YKRIPIPELQKADSPDFESPV-----MAGGLFAV 227
QY 265 TKDHLKKINGFSNDFWCGGEDDDLATRTSMAGLKVSRRP 304
DB 228 DRKMFELGSGYDGLIWMGEGQYISFKVS---QLSRRP 263
```

```
RESULT 11
US-11-134-241-39
Sequence 39, Application US/11/134241
Publication No. US20050287568A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Halbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kleke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
FILE REFERENCE: LEX-0144-USA
```

```
CURRENT APPLICATION NUMBER: US/11/134,241
CURRENT FILING DATE: 2005-05-20
PRIOR APPLICATION NUMBER: US/10/364,774
PRIOR FILING DATE: 2003-02-11
PRIOR APPLICATION NUMBER: US/09/795,926
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 366
TYPE: PRT
ORGANISM: homo sapiens
US-11-134-241-39
```

```
Query Match 3.9%; Score 80; DB 7; Length 366;
Best Local Similarity 17.9%; Pred. No. 2.5;
Matches 50; Conservative 41; Mismatches 75; Indels 114; Gaps 12;
```

```
QY 72 STFSPISEVNTQSFMEDIRPLFPDNTQLQFCNQTPHVLGPIRVFLDEDPFTLEK--- 128
DB 148 STIIFPHNEGWSLRTVHSV-----NRSPPELVAEIVLVDDFSDEHLKPLE 197
QY 129 ----IYP-----DTH--AGGHGMPK--DCVARR 148
DB 198 DYVALPFSVRILRTKKREGILRTRMGASVATGDTVITFLDSHCANVNMPLPILDRIRRN 257
QY 149 RVAIIVPYRDREAHRLIMLNHLSLAKQOLDVAIFVEQVANQTFNRGKLMNVGYDVAS 208
DB 258 RKTIVCPMD-----VIDHDDFRYETQA-----GDAMRGAPD--- 289
QY 209 RLVPMOCFTFHDVLLPE---DDRNLYTCPIQPRHMSVAIDKENYKLPYSALFGISAL 264
DB 290 ----WEMV-YKRIPIPELQKADSPDFESPV-----MAGGLFAV 324
QY 265 TKDHLKKINGFSNDFWCGGEDDDLATRTSMAGLKVSRRP 304
DB 325 DRKMFELGSGYDGLIWMGEGQYISFKVS---QLSRRP 360
```

```
RESULT 12
US-10-793-626-236
Sequence 236, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 236
LENGTH: 258
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-10-793-626-236
```

```
Query Match 3.8%; Score 78; DB 6; Length 258;
Best Local Similarity 23.7%; Pred. No. 2.4;
Matches 42; Conservative 24; Mismatches 47; Indels 64; Gaps 8;
```

```
QY 28 IYKIPSLYENLTIGSST----LIADVDAEAVLGNTASTSDLLDTWNSTFSPISERVNQ 82
```

Db 44 I E K I N L A D N I G I G H E T N K V L D T I F P R D M E M I A G N --- H D E A L --- M S L V N G 90
Qy 83 T S F M D I P I L E P D N O T I O --- F C N T P H L V G P I R --- V F --- 117
Db 91 T P Y P D L G K F E H H O W I E G H L D E S Y D E I N O L P R Y I E N T K K I L F H Y E I N D R K S A 150
Qy 118 - L D E D F K T L E K I Y P D T H A G H G M P D C V A R H R V A I I V P R D R E A H L R I M L H N - L H 171
Db 151 P I D E P F A P I T K --- D D E Q A I S E L F K X E A D L I I F G N H L H 189

RESULT 13
US-11-172-145-25
; Sequence 25, Application US/11172145
; Publication No. US20050272696A1
; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; PRIOR FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/491,362
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See file Wrapper or PALM.
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-25

Query Match 3.8%; Score 77.5; DB 7; Length 704;
Best Local Similarity 22.9%; Pred. No. 11;
Matches 30; Conservative 17; Mismatches 49; Indels 35; Gaps 5;

Db 171 H S L A K Q O L D V A I F I V E Q V A N O T E N R G K L M --- N G Y D V A S R L Y P M Q C F I H D V D L P E 226
Qy 275 H N I T A E Q F L N D P - Y I E S L P E T A T N N P S I T S K G N I S I D --- W R L E H F K K T D --- 322
Db 227 D D R N L Y T G P I O P R H M S V A I D K E N Y K L P Y S A I F G I S A L T K D L K K I N G S N D P M W G G E D 286
Qy 323 --- N U R L C D S P P R Y S C --- G N V A P S K E W L N K V G W P D E E F N H M G G E D 363
Db 287 D D L A T R T S M A G 297
Qy 364 V E F G Y R L P A K G 374

RESULT 14
US-11-172-145-4
; Sequence 4, Application US/11172145
; Publication No. US20050272696A1

; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; PRIOR FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/491,362
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See file Wrapper or PALM.
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-4

Query Match 3.8%; Score 77.5; DB 7; Length 965;
Best Local Similarity 22.9%; Pred. No. 17;
Matches 30; Conservative 17; Mismatches 49; Indels 35; Gaps 5;

Db 171 H S L A K Q O L D V A I F I V E Q V A N O T E N R G K L M --- N G Y D V A S R L Y P M Q C F I H D V D L P E 226
Qy 275 H N I T A E Q F L N D P - Y I E S L P E T A T N N P S I T S K G N I S I D --- W R L E H F K K T D --- 322
Db 227 D D R N L Y T G P I O P R H M S V A I D K E N Y K L P Y S A I F G I S A L T K D L K K I N G S N D P M W G G E D 286
Qy 323 --- N U R L C D S P P R Y S C --- G N V A P S K E W L N K V G W P D E E F N H M G G E D 363
Db 287 D D L A T R T S M A G 297
Qy 364 V E F G Y R L P A K G 374

RESULT 15
US-11-124-215-3
; Sequence 3, Application US/11124215
; Publication No. US20050266460A1
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTEURELLA MULTOCIDA AND
; FILE REFERENCE: 35541.101
; CURRENT APPLICATION NUMBER: US/11/124,215
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: 10/217,613
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 28

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:22:30 ; Search time 114 Seconds
(without alignments)
1403.760 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCANV.....VDLKEGCRRLRDPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US10a_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US10b_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	383	4	US-10-661-430-1 Sequence 1, App1
2	718	35.4	404	3	US-09-925-301-1072 Sequence 1072, Ap
3	710	35.0	393	4	US-10-132-652-5 Sequence 5, App1
4	710	35.0	393	6	US-11-105-796-5 Sequence 5, App1
5	709	34.9	403	6	US-11-097-143-18555 Sequence 18555, A
6	680.5	33.5	362	4	US-10-459-311-7 Sequence 7, App1
7	673.5	33.2	362	4	US-10-132-652-6 Sequence 6, App1
8	673.5	33.2	362	6	US-11-105-796-6 Sequence 6, App1
9	667.5	32.9	352	4	US-10-087-192-2049 Sequence 2049, Ap
10	666	32.8	385	4	US-10-466-941-6 Sequence 6, App1
11	666	32.8	385	5	US-10-870-635-6 Sequence 6, App1
12	666	32.8	398	5	US-10-761-435A-16 Sequence 16, App1
13	663.5	32.7	371	5	US-10-761-435A-20 Sequence 20, App1
14	660.5	32.5	397	5	US-10-713-970-21 Sequence 21, App1
15	660	32.5	372	4	US-10-287-226-82 Sequence 82, App1
16	657	32.3	372	6	US-10-132-652-2 Sequence 2, App1
17	657	32.3	372	6	US-11-105-796-2 Sequence 2, App1
18	649	32.0	400	4	US-10-132-652-4 Sequence 4, App1
19	649	32.0	400	6	US-11-105-796-4 Sequence 4, App1
20	646	31.8	342	4	US-10-132-652-3 Sequence 3, App1
21	646	31.8	344	6	US-11-105-796-3 Sequence 3, App1
22	642.5	31.6	344	3	US-09-989-722-236 Sequence 236, App
23	642.5	31.6	344	3	US-09-989-723-236 Sequence 236, App
24	642.5	31.6	344	3	US-09-989-727-236 Sequence 236, App
25	642.5	31.6	344	3	US-09-989-727-236 Sequence 236, App
26	642.5	31.6	344	3	US-09-989-731-236 Sequence 236, App
27	642.5	31.6	344	3	US-09-989-732-236 Sequence 236, App

28	642.5	31.6	344	3	US-09-991-073-236 Sequence 236, App
29	642.5	31.6	344	3	US-09-990-442-236 Sequence 236, App
30	642.5	31.6	344	3	US-09-991-163-236 Sequence 236, App
31	642.5	31.6	344	3	US-09-993-604-236 Sequence 236, App
32	642.5	31.6	344	3	US-09-990-456-236 Sequence 236, App
33	642.5	31.6	344	3	US-09-989-721-236 Sequence 236, App
34	642.5	31.6	344	3	US-09-992-598-236 Sequence 236, App
35	642.5	31.6	344	3	US-09-989-293A-236 Sequence 236, App
36	642.5	31.6	344	3	US-09-989-735-236 Sequence 236, App
37	642.5	31.6	344	3	US-09-990-444-236 Sequence 236, App
38	642.5	31.6	344	3	US-09-991-181-236 Sequence 236, App
39	642.5	31.6	344	3	US-09-989-730-236 Sequence 236, App
40	642.5	31.6	344	3	US-09-990-436-236 Sequence 236, App
41	642.5	31.6	344	3	US-09-993-687-236 Sequence 236, App
42	642.5	31.6	344	3	US-09-989-734-236 Sequence 236, App
43	642.5	31.6	344	3	US-09-997-653-236 Sequence 236, App
44	642.5	31.6	344	3	US-09-989-724-236 Sequence 236, App
45	642.5	31.6	344	3	US-09-989-728-236 Sequence 236, App

ALIGNMENTS

RESULT 1
US-10-661-430-1
; Sequence 1, Application US/10661430
; Publication No. US20040086955A1
; GENERAL INFORMATION:
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: Beta 1, 4-N-ACETYLGLACTOSAMINYLTRANSFERASES,
; FILE REFERENCE: 7148.001
; CURRENT APPLICATION NUMBER: US/10/661,430
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 60/411,242
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-661-430-1

Query Match	Best Local Similarity	Score	DB 4	Length	383
Best Match	100.0%	2031	DB 4	Length	383
Best Similarity	100.0%	Pred. No. 8.4e-196			
Matches	383	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	1	MAFRLAVARLKSILVLCANVLLVHAMTYKPSLYENTTIGSSTLIADVDAMEAVLGNTA	60		
DB	1	MAFRLAVARLKSILVLCANVLLVHAMTYKPSLYENTTIGSSTLIADVDAMEAVLGNTA	60		
QY	61	STSDLDLPTWNTSPFISVNOTSFMEDIRPLFPDNOTLOPCNOTPHLVGPIRVFIDE	120		
DB	61	STSDLDLPTWNTSPFISVNOTSFMEDIRPLFPDNOTLOPCNOTPHLVGPIRVFIDE	120		
QY	121	PDFKLEKTYPDYTAGGGMKDCVARRVAIIVYRREARLRIMLNHSLAKOOLD	180		
DB	121	PDFKLEKTYPDYTAGGGMKDCVARRVAIIVYRREARLRIMLNHSLAKOOLD	180		
QY	181	YAIFFVEQVANOTFNRGLAMVGVYASRLVPMOCFIFHDVDLPEDDRNLTCTPIOPRH	240		
DB	181	YAIFFVEQVANOTFNRGLAMVGVYASRLVPMOCFIFHDVDLPEDDRNLTCTPIOPRH	240		
QY	241	MSVAIDKENYKLPYSALFGISALTQDLKKINGFSNDPFWGSGEDDLATFTSMAGLKV	300		
DB	241	MSVAIDKENYKLPYSALFGISALTQDLKKINGFSNDPFWGSGEDDLATFTSMAGLKV	300		
QY	301	SRYPQIARVYKMGKSTATATPVKCRKYKMGOTGRRTGRLSTLYKLVNLEKPLVT	360		
DB	301	SRYPQIARVYKMGKSTATATPVKCRKYKMGOTGRRTGRLSTLYKLVNLEKPLVT	360		

QY 361 RAVVDLLEKDCRRLRRDPFTCF 383
|||
Db 361 RAVVDLLEKDCRRLRRDPFTCF 383

RESULT 2

US-09-925-301-1072
; Sequence 1072, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1072
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-301-1072

Query Match 35.4%; Score 718; DB 3; Length 404;
Best Local Similarity 48.9%; Pred. No. 2,9e-63;
Matches 134; Conservative 43; Mismatches 95; Indels 2; Gaps 2;

QY 95 PDNQTLOFCNQTPPHLVGPRIY-FLDEPDKTLEKIYPTDHAGHGMPKDCVARRHVAII 153
|||
Db 80 PAXQGLPYCERSPLLVGPVSVSPSLAEIYERPRVPEGRRYPACCEPRSRTAII 139
|||
QY 154 VPRYREAHRLIMLHNLSLAKQOLDYAFIYEOVANOTFNKGKLMNNGYDVASRLYPW 213
|||
Db 140 VPRAREHRLRLYLHHPFLQROQLAYGIYVHOGNGTFNPAKLLNVGREALRDEEM 199
|||
QY 214 QCFIFHDVDLPEDDENLYTC-PIOPRHMSVAIDKENYKLPYSALFGISALTCKDLKKI 272
|||
Db 200 DCLFLHDVDLPEDDENLYTC-PIOPRHMSVAIDKENYKLPYSALFGISALTCKDLKKI 259
|||
QY 273 NGFSNDFWGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 332
|||
Db 260 NGFPNEYWGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 319
|||
QY 333 QTKRWTRDGLSNLKYKLVNLEKPLYTRAIVDL 366
|||
Db 320 RTQNSWTQDGMNSLTVOQLARELGPLYTNTITADI 353
|||

RESULT 3
US-10-132-652-5
; Sequence 5, Application US/10132652
; Publication No. US20030013173A1

; GENERAL INFORMATION:
; APPLICANT: Claussen, Henrik
; APPLICANT: Beaumett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
; FILE REFERENCE: 4305/08521
; CURRENT APPLICATION NUMBER: US/10/132,652
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US/09/118,464A
; PRIOR FILING DATE: 1998-07-17

; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-132-652-5

Query Match 35.0%; Score 710; DB 4; Length 393;
Best Local Similarity 48.5%; Pred. No. 1.8e-62;
Matches 133; Conservative 43; Mismatches 96; Indels 2; Gaps 2;

QY 95 PDNQTLOFCNQTPPHLVGPRIY-FLDEPDKTLEKIYPTDHAGHGMPKDCVARRHVAII 153
|||
Db 69 PAXQGLPYCERSPLLVGPVSVSPSLAEIYERPRVPEGRRYPACCEPRSRTAII 128
|||
QY 154 VPRYREAHRLIMLHNLSLAKQOLDYAFIYEOVANOTFNKGKLMNNGYDVASRLYPW 213
|||
Db 129 VPRAREHRLRLYLHHPFLQROQLAYGIYVHOGNGTFNPAKLLNVGREALRDEEM 188
|||
QY 214 QCFIFHDVDLPEDDENLYTC-PIOPRHMSVAIDKENYKLPYSALFGISALTCKDLKKI 272
|||
Db 189 DCLFLHDVDLPEDDENLYTC-PIOPRHMSVAIDKENYKLPYSALFGISALTCKDLKKI 248
|||
QY 273 NGFSNDFWGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 332
|||
Db 249 NGFPNEYWGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 308
|||
QY 333 QTKRWTRDGLSNLKYKLVNLEKPLYTRAIVDL 366
|||
Db 309 RTQNSWTQDGMNSLTVOQLARELGPLYTNTITADI 342
|||

RESULT 4

US-11-105-796-5
; Sequence 5, Application US/11105796
; Publication No. US20050181437A1
; GENERAL INFORMATION:
; APPLICANT: CLAUSSEN, HENRIK
; APPLICANT: BEAUMETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; FILE REFERENCE: 04305/1008521-US1
; CURRENT APPLICATION NUMBER: US/11/105,796
; CURRENT FILING DATE: 2005-04-13
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-105-796-5

Query Match 35.0%; Score 710; DB 6; Length 393;
Best Local Similarity 48.5%; Pred. No. 1.8e-62;
Matches 133; Conservative 43; Mismatches 96; Indels 2; Gaps 2;

QY 95 PDNQTLOFCNQTPPHLVGPRIY-FLDEPDKTLEKIYPTDHAGHGMPKDCVARRHVAII 153
|||
Db 69 PAXQGLPYCERSPLLVGPVSVSPSLAEIYERPRVPEGRRYPACCEPRSRTAII 128
|||
QY 154 VPRYREAHRLIMLHNLSLAKQOLDYAFIYEOVANOTFNKGKLMNNGYDVASRLYPW 213
|||
Db 129 VPRAREHRLRLYLHHPFLQROQLAYGIYVHOGNGTFNPAKLLNVGREALRDEEM 188
|||
QY 214 QCFIFHDVDLPEDDENLYTC-PIOPRHMSVAIDKENYKLPYSALFGISALTCKDLKKI 272
|||
Db 189 DCLFLHDVDLPEDDENLYTC-PIOPRHMSVAIDKENYKLPYSALFGISALTCKDLKKI 248
|||
QY 273 NGFSNDFWGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 332
|||

Db 249 NGFENYMGWGEDDDIATRVRLAGMKISRPTSYGHYKMYVHGDKGNEENPHRFLLV 308
QY 333 QTKRRWRDGLSNLKYKLVNLEKPLYTRAVVDL 366
Db 309 RTQNSWTQDGMNSLTYQLLARELGPLYNTITADI 342

RESULT 5
US-11-097-143-18555

; Sequence 18555, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 18555
; LENGTH: 403
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-18555

Query Match 34.9%; Score 709; DB 6; Length 403;

Best Local Similarity 42.3%; Pred. No. 2.3e-62; Matches 150; Conservative 56; Mismatches 129; Indels 20; Gaps 7;

QY 28 IYKISLYENLT---IGSSTLIADVADMAVAGTASTSDDLDTWNSTF--SPISEVNOY 83
Db 46 VHKYAHITIGNASSDAGGSE-ASRLPASPALSCKRERQDELNGGPNSTIRIVIAIANFT 104
QY 84 SFMEDIRPL-----FPDNO-----TLQFCNQTPPHLVGPRIVFLDEDFKTLR-KIYP 131
Db 105 SIPOULTRFLGTGKFLPRKQSTSLANCTDPPRQDGPITPTTLESIDVIEAELGP 164
QY 132 DTHAGGCHKPKCVARHRAIIVPRDRRAHRIIMLHNSHLAQOQDYAIFYEQVAN 191
Db 165 LIRPGAGEPEPCNQHVAIVPFRDRYAHLLFLRNHPLMQRALVRIPIVEQTNQ 224
QY 192 QTFNRGLKLVNGVYDVAARLYPWCQFIHVDVLLPEDDRRLYTCPIOPRHMSVAIDKFNK 251
Db 225 KFPNDAAMNIGYLEALIKLYQWDCEFIHVDVLLPEDDRRLYTCPIOPRHMSVAIDKFNK 284
QY 252 LPEYSAIFGSIATYDHLKIKINGFSNDFWGWGEDDDIATRTSMAGLKVSRYPTQIARYK 311
Db 285 LPEYSAIFGSIATYDHLKIKINGFSNDFWGWGEDDDIATRTSMAGLKVSRYPTQIARYK 344
QY 312 MKHSTETANPNKCRKYKMGQTKRRTDGLSNLKYKLVNLEKPLYTRAVVDL 366
Db 345 MKHQKEKANPK---RYENLQNGMSKIEODGINSIKYSIYIKOPFTFTWYLAEL 396

RESULT 6
US-10-459-311-7

; Sequence 7, Application US/10459311
; Publication No. US2003020396A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Corley, Neil C.
; Shah, Purvi
; Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/459,311
; FILING DATE: 10-Jun-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/373,902
; FILING DATE: 12-Aug-1999
; APPLICATION NUMBER: US/09/055,097
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Ceirone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1463908
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-459-311-7

Query Match 33.5%; Score 680.5; DB 4; Length 362;

Best Local Similarity 43.9%; Pred. No. 1.5e-59; Matches 122; Conservative 58; Mismatches 97; Indels 1; Gaps 1;

QY 90 RPILEPDNQTLOFCNQTPPHLVGPRIVFLDER-DKTELEKIYPTDTHAGGCHKPKCVARH 148
Db 80 QPRNLPDAPSGLCDPSPPLLVGPRIVGFESQPVNLEAVASTNPEVEEGRFAPKCKKALQ 139
QY 149 RVAIVPRDRDEAHRIIMLHNSHLAQOQDYAIFYEQVANQTFNRGLKLVNGVYDVA 208
Db 140 KVAIIIFPRNREHKLKYLWYHPILOQOQDYGVVYINQDDEFFNRAKLINVFTEAL 199
QY 209 RLYPWCQFIHVDVLLPEDDRRLYTCPIOPRHMSVAIDKFNKLPDYSAIFGSIATYDHL 268
Db 200 KEYDYDCEVDFVDVLLIPMDRRTYKCYGQPRHLISYMDKFGFRLPYNOYFGGVSALSSEQ 259
QY 269 LKINGFENDFWGWGEDDDIATRTSMAGLKVSRYPTQIARYKMKHSTETANPNKCRKY 328
Db 260 FTKINGFENDFWGWGEDDDIATRTSMAGLKVSRYPTQIARYKMKHSTETANPNKCRKY 319
QY 329 KIMGQTKRRTDGLSNLKYKLVNLEKPLYTRAVVDL 366

RESULT 10
US-10-466-941-6
; Sequence 6, Application US/10466941
; Publication No. US20040214273A1
; GENERAL INFORMATION:
; APPLICANT: Tatsuji, Seki
; APPLICANT: Kazuhito, Fujiyama
; TITLE OF INVENTION: Method for excretory production of glycoproteins
; TITLE OF INVENTION: having human-type sugar chains
; FILE REFERENCE: J198020401
; CURRENT APPLICATION NUMBER: US/10/466,941
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: JP200112519
; PRIOR FILING DATE: 2001-01-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-466-941-6

Query Match 32.8%; Score 666; DB 4; Length 385;
Best Local Similarity 35.7%; Pred. No. 4,7e-58;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHAMLY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA----- 60
13 LVAVCALHLGTVLVYLAGRDSLRLPOL-----VGSTPLQGGNSAAAIQSSGSELRTG 67
61 -----STDDLDPWNSTSPISFVNQTSFEMDIRPLIFPDNQTLPQC 103
68 GARPPPLGASQPRGGSSPVVDSCGPPASNLTSV-----PVPHTTALSIPAC 117
104 NOTPHLVGPTRVFLDEP-DFKTEKITYPDTHAGHPKDCVARRHVAIIVPYDREAH 162
118 PRESPLVGPMLIEFMPVDLELVAKONPNVMGGRVAPRDCVSPHKVAIIIPFNRQEH 177
163 LRLMHNHSLAKQQLDVAIFIVEQVANQFNRGKLNNGVDVASRLYPMOCFIHFDV 222
178 LKTYLYLHPVLOKQQLDGIYVINOAGDTIFNRKALNVGFQELKQDYTCFVPSVDV 237
223 LIPEDDRNLVTCPIOPRHMSVAIDKENYKLPYSALFGISALTQKHLKKNFSDPFMGW 282
238 LIPMDHNAVRCFSPRHISVAMDKGFSLPYVQYFGVSALSQOPLTINGFPNNYMGW 297
283 GGEEDDLATRTSMAGLKVSRYPTQIARYKMIKSTEATNPVNKKRYKIMGOTKRRWTRDG 342
298 GGEEDDIFNRVLFVRGMSISRPNAVVGRCMIRHSRDKKNEPQRFDRLAHTKETMLSDG 357
343 LSNLKYKLVNLEKLYTRAVIDL 366
358 LNSLTYQVLDVQRYPLTYQITVDI 381

RESULT 11
US-10-870-635-6
; Sequence 6, Application US/10870635
; Publication No. US20050143564A1
; GENERAL INFORMATION:
; APPLICANT: Tatsuji, Seki and Kazuhito, Fujiyama
; TITLE OF INVENTION: Method for manufacturing of glycoproteins
; TITLE OF INVENTION: having human-type glycosylation
; FILE REFERENCE: 62,7368
; CURRENT APPLICATION NUMBER: US/10/870,635
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 09/857,651
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 385

TYPE: PRT
ORGANISM: Homo sapiens
US-10-870-635-6

Query Match 32.8%; Score 666; DB 5; Length 385;
Best Local Similarity 35.7%; Pred. No. 4,7e-58;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHAMLY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA----- 60
13 LVAVCALHLGTVLVYLAGRDSLRLPOL-----VGSTPLQGGNSAAAIQSSGSELRTG 67
61 -----STDDLDPWNSTSPISFVNQTSFEMDIRPLIFPDNQTLPQC 103
68 GARPPPLGASQPRGGSSPVVDSCGPPASNLTSV-----PVPHTTALSIPAC 117
104 NOTPHLVGPTRVFLDEP-DFKTEKITYPDTHAGHPKDCVARRHVAIIVPYDREAH 162
118 PRESPLVGPMLIEFMPVDLELVAKONPNVMGGRVAPRDCVSPHKVAIIIPFNRQEH 177
163 LRLMHNHSLAKQQLDVAIFIVEQVANQFNRGKLNNGVDVASRLYPMOCFIHFDV 222
178 LKTYLYLHPVLOKQQLDGIYVINOAGDTIFNRKALNVGFQELKQDYTCFVPSVDV 237
223 LIPEDDRNLVTCPIOPRHMSVAIDKENYKLPYSALFGISALTQKHLKKNFSDPFMGW 282
238 LIPMDHNAVRCFSPRHISVAMDKGFSLPYVQYFGVSALSQOPLTINGFPNNYMGW 297
283 GGEEDDLATRTSMAGLKVSRYPTQIARYKMIKSTEATNPVNKKRYKIMGOTKRRWTRDG 342
298 GGEEDDIFNRVLFVRGMSISRPNAVVGRCMIRHSRDKKNEPQRFDRLAHTKETMLSDG 357
343 LSNLKYKLVNLEKLYTRAVIDL 366
358 LNSLTYQVLDVQRYPLTYQITVDI 381

RESULT 12
US-10-761-435A-16
; Sequence 16, Application US/10761435A
; Publication No. US20040241817A1
; GENERAL INFORMATION:
; APPLICANT: Umana, Pablo
; APPLICANT: Bruenker, Peter
; APPLICANT: Ferrara, Claudia
; APPLICANT: Suter, Tobias
; TITLE OF INVENTION: Fusion Constructs and Use of Same to Produce Antibodies with
; TITLE OF INVENTION: Increased Fc Receptor Binding Affinity and Effector Function
; FILE REFERENCE: 1975.0180003
; CURRENT APPLICATION NUMBER: US/10/761,435A
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US 60/441,307
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/491,254
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/495,142
; PRIOR FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: GalT amino acid sequence from pbluegalT
US-10-761-435A-16

Query Match 32.8%; Score 666; DB 5; Length 398;
Best Local Similarity 35.7%; Pred. No. 5e-58;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHAMLY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA----- 60

```
Db 26 LVAVALHGLVTLVYVYLAGRDLRLPOL-----VGVSTPLGGSSNSAAAIGSSGELRTG 80
Qy 61 -----STSDLDLDTMNSNFSPISSEVNOQSFMEIDIRPILFPNOTLQPC 103
Db 81 GARPPPLGASGPPRGDSSPVVDSSGPGPASNLTISV-----PVPHTALSLPAC 130
Qy 104 NOTPPIHVGPIRVFLDEP-DFKTLKLYPDTNAGGMPKDCVARRHVAIIIVRYDREAH 162
Db 131 PEESPLLVGPMLEFNMFPVDLELVAKQNPVVKGGRYAPRDCVSPHKVAIIIFPRRQEH 190
Qy 163 LRLMLNHLSLAKQQLDYAIFIVEQVANOTFNRGKLMNVGVASRLPVMQCFITHVD 222
Db 191 LKMYLTYLHPVLOROQDLDPYGVINQAGDTIFNRAKLLNVGFQELKDYDTCTCFVSDVD 250
Qy 223 LPEDEDRNIYTCPIOPRHMSVALDKENYKLPYSAIFGGSIALTKDHLKKNIGFSNPMGW 282
Db 251 LIPMNDHNAVRCGSPRHHSIVAMDKFGSLPYQYFGGVASLSKQOFLTNGFPNNYMGW 310
Qy 283 GGEDDLATRTSMAGLKVSRYPTQIARVKMIKISTEATNPNKCRKIMQTRRWTRDG 342
Db 311 GGEDDDIFNRLVFRGMSISRPNVAVGRCRMIRSRDKNEPNOFRDRIAHTKETMLSDG 370
Qy 343 LSNLTKLVNLLELKPITYRAVVDL 366
Db 371 LNSLTQVLDVQRYPLTYQTIVDI 394
```

RESULT 13

```
US-10-761-435A-20
; Sequence 20, Application US/10761435A
; Publication No. US20040241817A1
; GENERAL INFORMATION:
; APPLICANT: Umana, Pablo
; APPLICANT: Bruenker, Peter
; APPLICANT: Riera, Claudia
; APPLICANT: Suter, Tobias
; TITLE OF INVENTION: Fusion Constructs and Use of Same to Produce Antibodies with
; FILE REFERENCE: 1975_0180003
; CURRENT APPLICATION NUMBER: US/10/761,435A
; PRIOR FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US 60/441,307
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/491,254
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/495,142
; PRIOR FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of ManII-GalT
US-10-761-435A-20
```

```
Query Match 32.7%; Score 663.5; DB 5; Length 371;
Best Local Similarity 36.1%; Pred. No. 8e-58;
Matches 133; Conservative 75; Mismatches 129; Indels 31; Gaps 6;
```

```
Qy 16 VLCAVLLVHAMI-----YKPSLYENLTIGSSTLIAD-VDAMEAVLGNTASTSDLL 67
Db 14 IFCVIISLYIMLDRGHLDYPRNRRRSGSPQGQLSMLOEKIDLEHLLENNEIISIR 73
Qy 68 DTNNTSPISSEVNOQSFMEIDIRPILFPNOTLOP-----CNQTPPIHVGPIRVFLD 119
Db 74 D-----SVIMLSSEVED-----GPKSSQSNFSQAGSPACPEESPLLVGPMLEFN 119
Qy 120 EP-DFKTLEKLYPDTNAGGMPKDCVARRHVAIIIVRYDREAHRLIMLHLSLLAKQ 178
Db 120 MPVDLELVAKQNPVVKGGRYAPRDCVSPHKVAIIIFPRRQEHKLTWLYLHPVLOQ 179
```

```
Qy 179 LDYAIFIVEQVANOTFNRGKLMNVGVASRLYPVMQCFITHVDLIPEDDRNLVTCPIOP 238
Db 180 LDYGVYINQAGDTIFNRAKLLNVGFQELKDYDTCTCFVSDVDLIPMNDHNAVRCGSP 239
Qy 239 REMSVAILDENKRLVYSAIFGGSIALTKDHLKKNIGFSNDPMWGGEEDDLATRTSMAGL 298
Db 240 RHISVAMDFGSLPYQYFGGVASLSKQOFLTNGFPNNYMGWGEEDDLINRLVFRGM 299
Qy 299 KVSRYPTQIARVKMIKISTEATNPNKCRKIMQTRRWTRDGLSNLKYKLVNLELKP 358
Db 300 SISRPNAVVGRCRMIRHSRDKNEPNOFRDRIAHTKETMLSDGLNSLTQVLDVQRYVL 359
Qy 359 YTPRAVVDL 366
Db 360 YTOITVDI 367
```

RESULT 14

```
US-10-713-970-21
; Sequence 21, Application US/10713970
; Publication No. US2005010664A1
; GENERAL INFORMATION:
; APPLICANT: Contreras, Roland
; APPLICANT: Callewaert, Nico L. M.
; APPLICANT: Verwecken, Wouter
; APPLICANT: Kaigorodov, Vladimir
; TITLE OF INVENTION: MODIFICATION OF PROTEIN GLYCOSYLATION IN METHYLOTROPHIC YEAST
; FILE REFERENCE: 17106
; CURRENT APPLICATION NUMBER: US/10/713,970
; CURRENT FILING DATE: 2003-11-14
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-713-970-21
```

```
Query Match 32.5%; Score 660.5; DB 5; Length 397;
Best Local Similarity 35.2%; Pred. No. 1.8e-57;
Matches 138; Conservative 73; Mismatches 134; Indels 47; Gaps 5;
```

```
Qy 7 AVARLKSLLVLCVLLVHAMIY-----KIPSLYENLTIGSSTLIADVDAMEAVLGN 58
Db 17 SLQRAERLLVAVCVMHGLVTLVYVYLAGRDLRLPOL-----VGVSTPLGGSSNSAAIIG 71
Qy 59 TA-----STSDLDLDTMNSNFSPISSEVNOQSFMEIDIRPILFP 95
Db 72 SSSELRTGARPPPLGASGPPRGDSSPVVDSSGPGPASNLTISV-----PVPHT 121
Qy 96 DNQTLQFCNQTPPIHVGPIRVFLDEP-DFKTLKLYPDTNAGGMPKDCVARRHVAIIIV 154
Db 122 TALSLPACPEESPLLVGPMLEFNMFPVDLELVAKQNPVVKGGRYAPRDCVSPHKVAII 181
Qy 155 PYVDREAHRLIMLHLSLLAKQQLDYAIFIVEQVANOTFNRGKLMNVGVASRLYPWQ 214
Db 182 PFRNRQEHKLTWLYLHPVLOROQDLDPYGVINQAGDTIFNRAKLLNVGFQELKDYDT 241
Qy 215 CFIEFHVVDLPEDDRNLVTCPIOPRHMSVALDKENYKLPYSAIFGGSIALTKDHLKKNIG 274
Db 242 CFVPSVDLIPMNDHNAVRCGSPRHHSIVAMDKFGSLPYQYFGGVASLSKQOFLTNG 301
Qy 275 FSNDFMGWGEEDDLATRTSMAGLKVSRYPTQIARVKMIKISTEATNPNKCRKIMQTR 334
Db 302 FPNNYMGWGEEDDDIFNRLVFRGMSISRPNVAVGRCRMIRHSRDKNEPNOFRDRIAH 361
Qy 335 KRWTRDGLSNLKYKLVNLELKPITYRAVVDL 366
Db 362 KETMLSDGLNSLTQVLDVQRYPLTYQTIVDI 393
```

```
RESULT 15
US-10-287-226-82
```



```

Sequence 82, Application US/10287226
Publication No. US20040086875A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ferenc,
APPLICANT: Burgeess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eissen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangolli, Esna A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekkuda, Ramesh,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malysankar, Uriel M.,
APPLICANT: Macdougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Murailidhara,
APPLICANT: Patturajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vernhet, Corine A.M.,
APPLICANT: Zernhusen, Bryan D.,
APPLICANT: Zhong, Wei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/334,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: Curoseq1st version 0.1
SEQ ID NO 82
LENGTH: 372
TYPE: PRF
ORGANISM: Homo sapiens
US-10-287-226-82

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:15:29 / Search time 39 Seconds
(without alignments)
944.898 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRIHAVRLKSLVLCAV.....VDLLEKDCRRLRDPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	667.5	32.9	399	2	AJ3396	beta-N-acetylgluco
2	666	32.8	397	2	UQ1030	beta-N-acetylgluco
3	648.5	31.9	402	2	S05018	beta-N-acetylgluco
4	569.5	28.0	224	2	T46511	hypothetical prote
5	523.5	25.8	490	2	A55141	glcNAc beta-1,4-N-
6	487.5	24.0	370	2	T26091	hypothetical prote
7	259	12.8	289	2	S40716	galactosyltransfer
8	244.5	12.0	194	2	PC1175	beta-N-acetylgluco
9	114	5.6	118	2	A39567	thymidine-binding
10	107.5	5.3	1105	2	S76557	carbamoyl-phosphat
11	105.5	5.2	715	2	T25233	hypothetical prote
12	103	5.1	1103	2	H82884	multiple banded an
13	99.5	4.9	623	2	T42245	probable polypepti
14	99.5	4.9	624	2	T42247	polypeptide N-acet
15	99.5	4.9	626	2	T42246	polypeptide N-acet
16	98	4.8	1175	2	S39551	chitin synthase (B
17	96	4.7	579	2	T31549	polypeptide N-acet
18	94	4.6	368	2	S47312	cysteine protease
19	93.5	4.6	585	2	B69718	conserved hypotet
20	93.5	4.6	1187	2	F86422	carbamoyl-phosphat
21	93	4.6	1510	2	D64120	cell division prot
22	92.5	4.6	363	2	A83470	probable periplasm
23	90.5	4.5	620	2	F83169	secretion protein
24	90.5	4.5	876	1	A57989	secretion protein
25	90	4.4	431	2	C88957	regulatory protein
26	89.5	4.4	443	2	AF1901	hypothetical prote
27	89	4.4	563	2	A88515	polypeptide N-acet
28	89	4.4	612	2	T42243	probable polypepti
29	88.5	4.4	647	2	C87440	metallo-beta-lacta

30	88.5	4.4	1164	2	F72166	A25R protein - var
31	88	4.3	358	2	T29492	hypothetical prote
32	88	4.3	1785	2	S53976	probable membrane
33	87.5	4.3	276	2	AG3532	hypothetical prote
34	87.5	4.3	314	2	G90365	conserved hypotet
35	87.5	4.3	509	2	AE1295	Bitfunctional phosph
36	87.5	4.3	776	2	T15411	hypothetical prote
37	87.5	4.3	872	2	E71852	valine-tRNA ligase
38	87.5	4.3	1156	2	T37411	RNA polymerase sub
39	87.5	4.3	1164	1	RNY28T	DNA-directed RNA p
40	87.5	4.3	1164	1	T28566	DNA-directed RNA p
41	87.5	4.3	2083	2	T42721	CRP-ductin-alpha p
42	87	4.3	336	2	A97168	glycosyltransferase
43	87	4.3	661	2	S37590	beta-fructofuranos
44	87	4.3	736	2	T40904	pombe specific hyp
45	87	4.3	1215	2	S60904	hypothetical prote

ALIGNMENTS

RESULT 1
A33396
beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - mou
N:Alternate names: glycoprotein 4-beta-galactosyltransferase; UDP-beta-1,4-galactosyltr
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A33396; A28799; JX0041; I49531; I49532
R:Hollis, G.F.; Douglas, J.G.; Shaper, N.L.; Shaper, J.H.; Stafford-Hollis, J.M.; Evans,
Biochem. Biophys. Res. Commun. 162, 1069-1075, 1989
A:Title: Genomic structure of murine beta-1,4-galactosyltransferase.
A:Reference number: A33396; MUID:89350913; PMID:2504153
A:Accession: A33396
A:Molecule type: DNA
A:Residues: 1-359 <HOL>
A:Cross-references: UNIPROT:P15535; UNIPARC:UPI0000022B9E; GB:M27922; NID:G341810; PIDN:
A>Note: the authors did not translate the codon GTT for residue 139
R:Shaper, N.L.; Hollis, G.F.; Douglas, J.G.; Kirsch, I.R.; Shaper, J.H.
J. Biol. Chem. 263, 10420-10428, 1988
A:Title: Characterization of the full length cDNA for murine beta-1,4-galactosyltransfer
A:Reference number: A28799; MUID:88273147; PMID:3134348
A:Accession: A28799
A:Molecule type: mRNA
A:Residues: 1-359 <SHA>
A:Cross-references: UNIPARC:UPI0000022B9E; EMBL:J03880; NID:G192185; PIDN:AAA37297.1; P
A>Note: It is uncertain whether Met-1 or Met-14 is the initiator
R:Nakazawa, K.; Ando, T.; Kimura, T.; Narimatsu, H.
J. Biochem. 104, 165-168, 1988
A:Title: Cloning and sequencing of a full-length cDNA of mouse N-acetylglucosamine (beta)
A:Reference number: JX0041; MUID:89033997; PMID:3141392
A:Accession: JX0041
A:Molecule type: mRNA
A:Residues: 1-359 <NAK>
A:Cross-references: UNIPARC:UPI0000022B9E; GB:D00314; NID:G220340; PIDN:BA000216.1; PID:
R:Shaper, J.H.; Hollis, G.F.; Shaper, N.L.
Biochimie 70, 1683-1688, 1988
A:Title: Evidence for two forms of murine beta-1,4-galactosyltransferase based on cloning
A:Reference number: I49531; MUID:89207607; PMID:3149531
A:Accession: I49531
A:Molecule type: mRNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: I49531
A:Cross-references: UNIPARC:UPI000016CC32; GB:M36289; NID:G192188; PIDN:AAA37294.1; PID:
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 14-63 <RE2>
A:Cross-references: UNIPARC:UPI000000492; GB:M36289; NID:G192188; PIDN:AAA37295.1; PID:
C:Genome: 139/1; 217/3; 280/2; 321/2; 356/2
A:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase; transmembrane protein
F:25-44/Domain: transmembrane #status predicted <TM>
F:113/Binding site: carbohydrate (Aen) (covalent) #status predicted

	Query Match	32.9%	Score 667.5	DB 2	Length 399
	Best Local Similarity	45.7%	Pred. No. 3,1e-50		
	Matches 123	Conservative 52	Mismatches 93	Indels 1	Gaps 1
QY	99	TLQFCNQTPEHLVSGIRV-FLDEBDDPKTLEKIYPTDTHAGHGMPDCYARRHVAIIYPR	157		
DB	127	SLPACPEESPPLVAVPMILDFNIAYDLIELAKKNPEIKTGGYSRDCSPKAXHIIIPFR	186		
QY	158	DREAHRLRLMNLHSLAKQOOLDYAIFVEQVAQNTENRGKLMVGVYDVASRLTPMOCFI	217		
DB	187	NRGHKLKMWLYLHPILQROQDDYGIYVINGAGTMEFRALNLINGEALAKOYDYNCFV	246		
QY	218	FHDVDDLPEDDRNLVTCPIDPRHMSVAIDKNYKLPYSALITGGISALTKHLKKNFSN	277		
DB	247	FSDVDLIPMDRRNRYRCFSOPRHISVANDKFGFSLPYVQYGVGSALSCKOOFIAINEPN	306		
QY	278	DFMGMGEDDDLATRTSMAGLKVSRYPQIARLYMKIKISTEATNPVNCRCRYKIMGOTRR	337		
DB	307	NYMGMGEDDDIIFNRLVHKMSISRPNNAVGRCMIRHSRDKKNPEPNQRDRLAHKTET	366		
QY	338	WTRDGLSLWKYLVNLELKLPLYTRAVDUJ	366		
DB	367	MRFDGLSLVRYKVLVDQRYPLTYQITVDYI	395		

RESULT 2
JQ1030
beta-N-acetylglucosaminylglycopeptidase beta-1,4-galactosyltransferase (EC 2.4.1.38) - human
N/Alternate names: N-acetylglucosamide (beta-1-4)-galactosyltransferase; UDP-beta-1,4-galactose
N/Contents: lactose synthase (EC 2.4.1.22) protein A; N-acetylglucosamine synthase (EC 2.4.1.22)
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 09-Jul-2004
C/Accession: JQ1030; A31353; S13792; S37717; A24251; A24229; S12047
R/Mengle-Gaw, L.; McCoy-Hawman, M.F.; Tiemeier, D.C.
Biochem. Biophys. Res. Commun. 176, 1269-1276, 1991
A>Title: Genomic structure and expression of human beta-1,4-galactosyltransferase.
A/Accession: JQ1030; MUID:91248214; PMID:1905938
A/Molecule type: DNA
A/Status: 1-397 <MEN>
A/Cross-references: UNIPROT:P15291; UNIPARC:UPI000012671C; GB:IN70427; GB:MT0431
R/Mari, K.A.; Appert, H.E.; Fukuda, M.N.
Biochem. Biophys. Res. Commun. 157, 657-663, 1988
A>Title: Identification of the full-length coding sequence for human galactosyltransferase
A/Accession: A31353; MUID:89076299; PMID:3144273
A/Molecule type: mRNA
A/Status: 1-211, 'GIT', 212-397 <MAS>
A/Cross-references: UNIPARC:UPI0000158706; GB:M22921
R/Watzele, G.; Berger, E.G.
Nucleic Acids Res. 18, 7174, 1990
A>Title: Near identity of HeLa cell galactosyltransferase with the human placental enzyme
A/Reference number: S13792; MUID:91088335; PMID:2124683
A/Accession: S13792
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-10, 'S', 11-397 <WAT>
A/Cross-references: UNIPARC:UPI00002022E; EMBL:X55415; NID:G32057; PIDN:CAA39073.1; PIR
R/Uejima, T.; Uemura, M.; Nozawa, S.; Narimatsu, H.
Cancer Res. 52, 6158-6163, 1992
A>Title: Complementary DNA cloning for galactosyltransferase associated with tumor and
A/Reference number: S37717; MUID:93046112; PMID:1384956
A/Accession: S37717
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: mRNA
A/Residues: 1-9, 'RS', 11-33, 'R', 35-397 <UEJ>
A/Cross-references: UNIPARC:UPI0000158549; EMBL:X13223; NID:G34989; PIDN:CAA31611.1; PIR
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1988
R/Appert, H.E.; Rutherford, T.J.; Tarr, G.E.; Wiest, J.S.; Thomford, N.R.; McCorquodale
Biochem. Biophys. Res. Commun. 139, 163-166, 1986
A>Title: Isolation of a cDNA coding for human galactosyltransferase.
A/Reference number: A24251; MUID:87025694; PMID:3094506

A:Accession: A24251
A:Molecule type: mRNA
A:Residues: 137-397 <APP>
A:Cross-references: UNIPARC:UPI000016A9F4; GB:M13701; MID:g183659; PIDN:AAA35935.1; PID:g183659
R:Appert, H.E.; Rutheford, T.J.; Tarr, G.E.; Thomford, N.R.; McCosquedale, D.J.
Biochem. Biophys. Res. Commun. 138, 224-229, 1986
A:Reference number: A24229; MUID:86295672; PMID:3091013
A:Accession: A24229
A:Molecule type: protein
A:Residues: 78-93 <AP2>
A:Cross-references: UNIPARC:UPI0000179825
A:Experimental source: milk
R:Aoki, D.; Appert, H.E.; Johnson, D.; Wong, S.S.; Fukuda, M.N.
EMBO J. 9, 3171-3178, 1990
A:Title: Analysis of the substrate binding sites of human galactosyltransferase by proteolysis
A:Reference number: S12047; MUID:91006010; PMID:2120039
A:Accession: S12047
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 273-325 <AOK>
A:Cross-references: UNIPARC:UPI0000179826
C:Comment: This enzyme catalyzes the transfer of galactose from an activated UDP-galactose to a protein
C:Genetics:
A:Gene: GDB:GCTB2
A:Cross-references: GDB:119981; OMIM:137060
A:Map position: 9p21-9p13
A:Introns: 136/3; 215/3; 278/3; 319/3; 355/3
C:Keywords: glycosyltransferase; hexosyltransferase; lactose biosynthesis; transmembrane

```

Query Match Similarity 32.8%; Score 666; DB 2; length 397;
Best local similarity 35.7%; Pred. No. 4, 1e-50;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

QY 14 LVLCAVLLVHAMTY-----KIPSLYENLTIGSSTLIADVADAEAVLGNTA----- 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 25 LVAVCALHGVTLVYVYLAGRDLRLPOL-----VGVSITPLOGGSNSAAAGSGSELRTG 79
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 -----STDDLLDTWNTSNFSPISSEVNGTCSFMEDIRPIPLPDDQTQLOC 103
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 80 GARPPPLGASQPRRGSGSSPVDSGPGPSNLTV-----PVHTTALSLPAC 129
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 104 NOTPHLVGVIRFLDEP-DFKTLLEKIYDPTNAGHGHPKDCVARRHVAIIIVYRDEAH 162
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 130 PESEPLVGVPMILIEFMFPVDLELVAKONPNVKGGRYAPRDCVS PHEVAIIIPFRNQH 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 LRIIMHNHSLIAKQOOLDYAIFVEOVANOTFNRGKLMNVGVVASRLTYMOCFIPIHDD 222
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 LKMYLYLHPVLQROQLDYGIVINQASDPIIFRAKLINVGFQELKDYTCVFSDDVD 249
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 LIPEDDRMLYTCPIOPRHSNSVAIDKFNYSALIPYSALIFGIGSALTQHLKKINGFSNDPWG 282
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 250 LIPMDHNHAYRCFSGOPRHISVAMDKGFSLPYQVQFGVSALSKOQPLTINGFPNNWGW 309
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 GGEDDDLARTSMAGIKVSRYPQTQIARVYKMIKSTEAETNVNKKCRXYKMGOTKKRMTRDG 342
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 310 GGEDDDIFRKLVRFGKSIIRPNNAVVGRCRMTIRSRDKGKEPNQRFPRIRATKETMLSDG 369
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 343 LSNLKYKLVNLELKLPIYRAVVDL 366
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 370 INSLTYQVADVORPYLYTOITVDI 393
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
S05018
beta-N-acetylglucosaminylglycopeptidase beta-1,4-galactosyltransferase (EC 2.4.1.38)
N/Alternative names: beta 1,4-galactosyltransferase
N/Contains: beta-N-acetylglucosaminylglycopeptidase beta-1,4-galactosyltransferase (EC 2.4.1.38)
C/Species: Bos primigenius taurus (cattle)
C/date: 07-Sep-1990 #sequence revision 07-Sep-1990 #text change 09-Jul-2004
C/accession: S05018; A35077; A24148; A25129; A39076; A37809; B35077
R/d/Agostaro, G.; Bendiak, B.; Tropak, M.
Eur. J. Biochem. 189, 211-217, 1989
AltTitle: Cloning of cDNA encoding the membrane-bound form of bovine beta 1,4-galactosyltr

```

A:Reference number: S05018; MUID:89325338; PMID:2502398
A:Molecule type: mRNA
A:Residues: 1-402 <DAG>
A:Cross-references: UNIPROT:P08037; UNIPARC:UPI0000168534; EMBL:X14558; NID:g381; PIDN:C
A:Experimental source: liver
A>Note: part of this sequence was confirmed by protein sequencing
R:Russo, R.N.; Shaper, N.L.; Shaper, J.H.
J. Biol. Chem. 265: 3324-3331, 1990
A:Title: Bovine beta1->4-galactosyltransferase: two sets of mRNA transcripts encode two
oath the short and the long forms of the enzyme are type II membrane-bound glycoproteins
A:Reference number: A35077; MUID:90153986; PMID:2105947
A:Accession: A35077
A:Molecule type: mRNA
A:Residues: 1-157,'V',159-165 <RUS>
A:Cross-references: UNIPARC:UPI000016C319; GB:J05217; NID:g16125; PIDN:AAA30559.1; PID:
R:Narimatsu, H.; Sinha, S.; Brew, K.; Okayama, H.; Qasba, P.K.
Proc. Natl. Acad. Sci. U.S.A. 83, 4720-4724, 1986
A:Title: Acid. Sci. U.S.A. 83, 4720-4724, 1986
A:Reference number: A24148; MUID:86259693; PMID:3014508
A:Accession: A24148
A:Molecule type: mRNA
A:Residues: 74-157,'V',159-163,'E',165-186,'P',188-204,'IL',207-255,257-264,'O',266-402
A:Cross-references: UNIPARC:UPI0000179827; GB:M13569; NID:g163407; PIDN:AAA30659.1; PID:
A:Experimental source: mammary gland
R:Shaper, N.L.; Shaper, J.H.; Neuch, J.L.; Fox, J.L.; Chang, H.; Kirsch, I.R.; Hollis, G
Proc. Natl. Acad. Sci. U.S.A. 83, 1573-1577, 1986
A:Title: Bovine galactosyltransferase: identification of a clone by direct immunological
A:Reference number: A25129; MUID:86149345; PMID:2419911
A:Accession: A25129
A:Molecule type: mRNA
A:Residues: 69-157,'V',159-186,'P',188-204,'IL',207-281,'V',283-402 <SHA>
A:Cross-references: UNIPARC:UPI000016C30F; GB:M13224; NID:g163073; PIDN:AAA30534.1; PID:
A:Experimental source: MDBK cells
R:Yadav, S.P.; Brew, K.
J. Biol. Chem. 266, 698-703, 1991
A:Title: Structure and function in galactosyltransferase. Sequence locations of alpha-1a
A:Reference number: A39076; MUID:91093259; PMID:1898734
A:Accession: A39076
A:Molecule type: protein
A:Residues: 81-95,'X',91-95;171-180;190-198,'YL',201;209-216;229-238,'H',240-244;351-358
A:Cross-references: UNIPARC:UPI0000179828; UNIPARC:UPI0000179829; UNIPARC:UPI000017982A;
82F; UNIPARC:UPI0000179830
R:Yadav, S.; Brew, K.
J. Biol. Chem. 265, 14163-14169, 1990
A:Title: Identification of a region of UDP-galactose-N-acetylglucosamine beta4-galactosy
A:Reference number: A37809; MUID:90354395; PMID:2117606
A:Accession: A37809
A:Molecule type: protein
A:Residues: 81-89,'X',91-95;171-180;190-198,'YL',201;209-216;229-238,'H',240-244;351-358
A:Cross-references: UNIPARC:UPI0000179831; UNIPARC:UPI0000179832; UNIPARC:UPI0000179833;
C:Genetics:
A:Gene: galT
C:Keywords: alternative initiators; glycoprotein; glycosyltransferase; hexosyltransferase
F:1-18/Domains: signal sequence #status predicted <SIG>
F:14-402/Product: beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase,
F:15-402/Product: beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase,
F:25-44/Domains: transmembrane #status predicted <TM>
F:729-250/Region: alpha-lactalbumin binding #status predicted
F:341-351/Region: UDP-galactose binding #status predicted
F:90,117/Binding site: carbonylate (asn) (covalent) #status experimental
F:134-247/Disulfide Bonds: #status experimental

Db 182 VAILIFPNRQEHKLYWLYLHPWQROQLDGIYVINGAGSEMFRAKLINVGFEKALK 241

Qy 210 LYPWOCFIHFHVDVLLPEDDRNLVTCPIQPRHMSVAIDKFNKYKLPYSALFGISALTQDL 269

Db 242 DYDNCFFFSVDVLLPMDNDHNTYRCFSQPRHISVAMDKFGSLPYQYFGVSAISKOQF 301

Qy 270 KKNESFSDPFWMGGEEDDLATRTSMAGLKVSRRPYQIARVYKMIKHSTEAINTVNCRYK 329

Db 302 LSNFPPNNYWMGGEEDDIYNRLAFRGMSVSRPNAVICKRMTIRHSRDKNEPNPQF 361

Qy 330 IMGQTRRTRDGLSNLKYLVLNLELPLRYTAVDL 366

Db 362 RIATKRTMLSDGLNSLTMTVLVQRYPLTYITVDI 398

RESULT 4

T46511

hypothetical protein DKFZp586M2424.1 - human (fragment)

C.Species: Homo sapiens (man)

C.Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 09-Jul-2004

C.Accession: T46511

R.Dusterhoef, A.J. Lauber, J.J. Mewes, H.W. Weil, B.J. Wiemann, S. submitted to the Protein Sequence Database, January 2000

A.Reference number: Z23036

A.Accession: T46511

A.Statut: preliminary

A.Molecule type: mRNA

A.Residues: 1-224 <AAA>

A.Cross-references: UNIPROT:060909; UNIPARC:UP1000016ACC; EMBL:AL137647

A.Experimental source: adult uterus, clone DKFZp586M2424

C.Genetics:

A.Note: DKFZp586M2424.1

Query Match 28.0%; Score 569.5; DB 2; Length 224;

Best Local Similarity 46.1%; Pred. No. 4,8e-42;

Matches 101; Conservative 44; Mismatches 63; Indels 11; Gaps 2;

Qy 154 VPYRDEAHLRIHNLHSLAKQOLDVAIFIVEQVANOTFNRGKIMNGY-----DVA 207

Db 1 IPRFRHRLRLWTLHYLHPIILRQRLRYGVYVINGEDTFFNRKLLNVGFLEALKEDAA 60

Qy 208 SRLVWOCFIHFHVDVLLPEDDRNLVTCPIQPRHMSVAIDKFNKYKLPYSALFGISALTQD 267

Db 61 ----YDFIFESDVLVPMDDENLYRCGDPHFALMDKFGFRLVAGYFGVSGISLKA 115

Qy 268 HLKKNINGSNDPMWGGEEDDLATRTSMAGLKVSRRPYQIARVYKMIKHSTEAINTVNCRYK 327

Db 116 QPLRNGFPNRYMGGEEDDLFNRISLTGMKISRDPDITIGRIRMLKHDRDKINNEPQR 175

Qy 328 YKIMGQTRRTRDGLSNLKYLVLNLELPLRYTAVDL 366

Db 176 FTKIQNTKLTMKRDGIGSVRYQVLEVSROPFLFTNIVDI 214

RESULT 5

A55141

glnac beta-1,4-N-acetylglucosaminyltransferase (EC 2.4.1.-) - great pond snail

C.Species: Lymnaea stagnalis (great pond snail)

C.Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #ext_change 09-Jul-2004

C.Accession: A55141

R.Bakker, H.J. Agteleberg, M.J. Van Tetering, A.J. Koelmaan, C.A.M. Van den Bijnden, D.H. J. Biol. Chem. 269, 30326-30333, 1994

A.Title: A Lymnaea stagnalis gene, with sequence similarity to that of mammalian beta-1,4-N-acetylglucosaminyltransferase

A.Reference number: A55141; MUID:95074032; PMID:7527028

A.Accession: A55141

A.Statut: preliminary; not compared with conceptual translation

A.Molecule type: mRNA

A.Residues: 1-490 <BAK>

A.Cross-references: UNIPROT:009323; UNIPARC:UP1000012676B; GB:X80228; NID:G620087; PID:G620087

C.Keywords: glycosyltransferase; hexosyltransferase

Query Match 25.8%; Score 523.5; DB 2; Length 490;

	Best Local Similarity 30.6%; Pred. No.1.5e-37; Matches 131; Conservative 66; Mismatches 158; Indels 73; Gaps 8;
QY	4 RHLAVARIKSLIVTCAVLLLVHAMIYKIPSLYENL-----TIGSSTLLADVAME 53
Db	27 RSVAVIKATAMLFVAAMLFALAHNMFSEHAQQONLHRAPISSPTTISRSTVQIRNATHD 86
QY	54 AVLQNTASTDDLLDTNNSTF-----SPISEVQTS--FMEDIRPLIFPNQTLQACNQ 105
Db	87 FLPASSTPMKQELLET--ESEFVDGQRENEVIACSDTSSEFTSDKRTLLVNSQSGVCPPI 145
QY	106 TPPIHVGIRVFLEDPDEPKTEKIYPRTHAGGHGPKDCVARRHRAVAILVPRYDEAHARI 165
Db	146 RPPALAGFVFSKSSSTYHSLAAMPDVQDSGHTPRNCTPAEKTAIIIPRRNCRHLYT 205
QY	166 MLNHLISLAAQQLDYAIFIVEQVANQOTPRNGKLMNVGYDVASLLYPMQCTPHDVLLP 225
Db	206 LLPNLIPMLAMQNVDFTFIEVIEQETTPETPNKGILFNAGYIEALKYDNYDCFIHDVMIIP 265
QY	226 EDDRLLYTC-PIQPHMNSVAIDKFNYYKLPYBAIRGGISALPKDHLKKINGNSNDFMGWG 284
Db	266 IDDRMYVNCNKGQPVHFPBGVKKFYKLFYGGLFEGVGVFIREOPRLINGASNLVFGWG 325
QY	285 EDDDLATRTSMAGLKVRSRYPTQIARKKMIKISTEA--TNP----- 322
Db	326 EDDDLRDAVIMKLPFLKRTLAHGLYDMVSH-VELGAVNNVHSHKGAHSLYMLKALGVQ 384
QY	323 -----VNKCRYYKIMGOTKKRMTREDGLSNLKYYKLVN 352
Db	385 AGMNVHPNSKMPRLRFLDSVNHAFAEGAGMNVNPDPRFKIYSTSRORHVDGINSLYVNVTV 444
QY	353 LELKPLYT 360
Db	445 YRTSPLYT 452

Db 171 QTHLTRLIDFLIPILQRRRLDPRFIVTEQYQNDLFNKGRIINAAFIAESE--GVDCVVEH 229
 QY 220 DVDLLPEDNRNLVYTCPIOPRHKMSVALDKENYKLPYSALIPGGSALTCKHLKINGRSNDP 279
 Db 230 DVDMEFPQDDRNPVSCPPGPRHIGAFVSNIGYQLMWKEIYGVALVMSADYRAVNGYSNOF 289
 QY 280 MGVGGEDDDLLTRTSMAGLKVSRYPTQIARVYKMGH--STEATNPVVKCRKIKMGOTKRRW 338
 Db 290 WAMGGEDDDMGCRILISLNTTIERPNPETGRYSMLKHVRKRTAP--KLIYKLLGNANRV 347
 QY 339 TRDGLS 344
 Db 348 AYDGLN 353
 RESULT 7
 S40716
 galactosyltransferase homolog R10E11.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S40716
 R:Almscough, R.; Mortimore, B.
 submitted to the EMBL Data Library, December 1993
 A:Reference number: S40713
 A:Accession: S40716
 A:Molecule type: DNA
 A:Residues: 1-289 <AIN>
 A:Cross-references: UNIPROT:P34548; UNIPARC:UPI0000135EA7; EMBL:Z29095; NID:G436453; PID:
 C:Genetics:
 A:Introns: 41/2; 75/2; 104/3; 196/2; 246/3
 Query Match 12.8%; Score 259; DB 2; Length 289;
 Best Local Similarity 31.7%; Pred. No. 8, 1e-15;

RESULT 6
T26091
hypothetical protein W02B12.11 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26091
R:Swindburne, J.; Ainscough, R.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z20149
A:Accession: T26091
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-370 <MIL>
A:Cross-references: UNIPROT:Q21608; UNIPARC:UP1000017BC11; EMBL:Z66521; PIDN:CAA91401.1
A:Experimental source: clone W02B12
C:Genetics:
A:Gene: CESP:W02B12.11
A:Map position: 2
A:introns: 41/2; 102/2; 199/3; 336/2

[illegible]

	Query Match	24.0%	Score 487.5;	DB 2;	Length 370;
	Best Local Similarity	30.1%;	Pred. NO. 1.4e-34;		
	Matches	110;	Mismatches 132;	Indels 59;	Gaps 10;
QY	12 KSLVLCAYLLLVHAMTYKIPSLYEHLTGSS---TLADVDAMEAVLGNASTSDDLLD	68			
Dd	14 KLIIIFVAASLTILYYMLLNSTLNDERRIAAVEEQVSNNVVNQSGNSGVARHDIKD	73			
QY	69 TWNSTFSPISSEVNQSFMEDIRPLLPDPNQTLOFCNQ----TPPHVGFIKFLEDDPDK	124			
Dd	74 -----PVE-----LEIAR-----QLTFLFNMEBAVKPK--PPVKAIEDGSCP	110			
QY	125 TLEK-----ITPDTHAGGHGMKDCVARRRVALIVPYBDR	159			
Dd	111 IVERIPDLOGALPQAATLIQNQGEGVAIHHELPGSGSKRDDCQARDKIATIVIPYER	170			
QY	160 EAHRLIMNLHSLSLAQQDLVAPIFEQVANQTFNRGLKNMGDVASRLYPQCFTFH	219			

C:Species: Gallus gallus (chicken)
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: PC1175
R:Ghosh, S.; Basu, S.S.; Basu, S.
R:Biochem. Biophys. Res. Commun. 189, 1215-1222, 1992
A>Title: Isolation of a cDNA clone for beta1-4 galactosyltransferase from embryonic chick
A:Reference number: PC1175; MUID:93112014; PMID:1281993
A:Accession: PC1175
A:Molecule type: mRNA
A:Residues: 1-194 <GHO>
A:Cross-references: UNIPROT:Q7LZG0; UNIPARC:UPI0000179838; GB:112565; NID:9289793
A:Note: the reading frame extends between four stop codons shown as X for residues 42, 43
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match	12.0 %;	Score 244.5;	DB 2;	Length 194;
Best Local Similarity	30.9 %;	Pred. No. 8.5e-14;		
Matches 58;	Conservative 27;	Mismatches 68;	Indels 35;	Gaps 6;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:11:03 ; Search time 133 Seconds
(without alignments)
1265.279 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRIHVARLKSILVLCAV.....VDLLEKDCRELRDPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_21.*
2: geneseqp1980s.*
3: geneseqp1980s.*
4: geneseqp2000s.*
5: geneseqp2000s.*
6: geneseqp2003ds.*
7: geneseqp2003ds.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	383	8	ADM32968
2	718	35.4	404	3	AA843627
3	717	35.3	393	8	ABM80189
4	710	35.0	393	3	ABE85868
5	709	34.9	403	4	ABE83921
6	680.5	33.5	362	7	ADG63356
7	680.5	33.5	362	7	ADL27179
8	673.5	33.2	362	9	ABE85869
9	668.5	32.9	264	3	ABE85869
10	668.5	32.9	767	2	AAE5709
11	668	32.9	398	2	AAE5709
12	667.5	32.9	352	7	ABM85878
13	667.5	32.9	399	8	ADQ79982
14	667.5	32.9	399	8	ADQ79982
15	667.5	32.9	399	9	AEA55054
16	667.5	32.9	399	9	AEA51684
17	666	32.8	385	3	ABE83647
18	666	32.8	385	5	ABE83647
19	666	32.8	398	2	AAE88388
20	666	32.8	398	7	ADP18448
21	666	32.8	398	7	ADP18448
22	666	32.8	398	8	ADQ78358
23	666	32.8	398	8	ADQ78358
24	665.5	32.8	767	2	AAE5708

25	665	32.7	398	2	AAE5706	AAE5706	Galactose
26	663.5	32.7	371	8	ADQ91374	ADQ91374	Manit-Gal
27	661	32.5	475	7	ADP18451	ADP18451	N-acetyl
28	661	32.5	500	7	ADP18450	ADP18450	N-acetyl
29	660.5	32.5	565	4	AAU01864	AAU01864	MAN_ATGT
30	660.5	32.5	397	9	ADG69855	ADG69855	Human Gal
31	660.5	32.5	402	4	AAU01865	AAU01865	MAN285 bo
32	660.5	32.5	406	4	AAU01867	AAU01867	GNT galac
33	660.5	32.5	433	4	AAU01866	AAU01866	MAN GMGT
34	660	32.5	372	7	ADP94854	ADP94854	NOV
35	657	32.3	372	6	ABU08239	ABU08239	Human Bet
36	657	32.3	372	7	AAE39165	AAE39165	Human Bet
37	657	32.3	372	8	ABM80077	ABM80077	Tumour-as
38	657	32.3	372	8	ABE95865	ABE95865	Human bet
39	656.5	32.3	720	8	ADQ78360	ADQ78360	alpha-1,3
40	655.5	32.3	355	7	ADD68135	ADD68135	Human gal
41	655.5	32.3	383	7	ADD68087	ADD68087	Human gal
42	655.5	32.3	383	7	ADD68141	ADD68141	Human gal
43	655	32.3	398	8	ADQ79981	ADQ79981	Human bet
44	655	32.3	398	9	AEA51683	AEA51683	Human bet
45	655	32.3	402	2	AAE05933	AAE05933	Mouse bet

ALIGNMENTS

RESULT 1		
ID	ADM32968	Standard; protein; 383 AA.
XX	ADM32968;	
AC	ADM32968;	
XX	17-JUN-2004	(first entry)
DT	17-JUN-2004	(first entry)
XX	Amino acid sequence of a beta1,4-N-acetylglucosaminyltransferase.	
DE	beta1, 4-N-acetylglucosaminyltransferase; beta4GalNAcT; animal cell	
XX	glycoprotein; lactinac motif; IDN motif; GalNAcbeta4GlcNAc-R;	
KW	glycoprotein hormone; pituitary gland; glycoconjugate; parasite;	
KW	immune response; infection; enzyme.	
XX	Caenorhabditis elegans.	
OS	Caenorhabditis elegans.	
XX		
PH	Key	Location/Qualifiers
FT	Domain	7..29
FT	Modified-site	/note = putative transmembrane domain
FT	Region	37
FT	Region	/note= "N-glycosylation site"
FT	Modified-site	48..50
FT	Modified-site	/note= "DVD motif"
FT	Modified-site	71
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	81
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	97
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	104
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	191
FT	Region	/note= "N-glycosylation site"
FT	Region	220..222
FT	Region	/note= "DVD motif"
XX		
PN	WO2004024938-A2.	
XX	25-MAR-2004.	
PD	12-SEP-2003; 2003WO-US028833.	
XX	13-SEP-2002; 2002US-0411242P.	
PF	(CUMM/) CUMMINGS R D.	
XX	(KAWA/) KAWAR Z.	
XX		
PA		

Dy 333 QTKRRNRTRDGF.SNLKYYKL.VNLELKP.LYTRAVVDL 366
::|||::|||::|||::|||::|||::|||
Db 320 RTQNSWTQDGMSLT.YQLARELGPLYTNITAD I 353

RESULT 3
ABM80189
ID ABM80189 standard; protein; 393 AA

DE	Tumour-associated antigenic target (TAT) polypeptide PRO80689, SEQ:469
XX	
KM	Tumour-associated antigenic target; TAT, human; overexpression; cancer
KM	tumour; diagnosis; cell proliferative disorder; breast cancer;
KM	colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KM	central nervous system cancer; bladder cancer; pancreatic cancer;
KM	cervical cancer; melanoma; leukaemia; hybridisation probe;
KM	chromosome identification; chromosome mapping; gene mapping;
KM	gene therapy; cytostatic.

OS	Homo sapiens.
XX	
PN	W02004030615-A2.

PF	29-SEP-2003; 2003WO-US028547.
XX	
PR	02-OCT-2002; 2002US-0414971P.

PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN37523.

PT New tumor-associated antigenic target polypeptides and nucleic acids
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

PS Claim 12; SEQ ID NO 469; 7273pp; English.

The invention relates to human tumour-associated antigenic target (TAT) polypeptides, and their related nucleic acids. The TAT polypeptides are overexpressed in cancer tissues compared to normal tissues, and may thus serve as effective targets for the diagnosis and treatment of cancer in mammals. The invention also relates to nucleic acid and polypeptide sequences at least 80% identical to the TAT nucleic acids and polypeptides; expression vectors and host cells comprising a TAT nucleic acid; an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAT polypeptide of the invention.

Query Match	35.3%;	Score 717;	DB 8;	Length 393;
Best Local Similarity	48.9%;	Pred. No. 2.1e-67;		

Matches 134; Conservative 43; Mismatches 95; Indels 2; Gaps 2;

Qy	95	EDMOTLOFCNCPRIHVGIRY-FIDDEDPFKTEKYTDTHAGHMPKOCVAHRVAII	15.93
Ds	69	PABQGLPYCPBSPFLLVGVSVSPVSLAEIYERNRPVBBGRYRYPACCEPSRPAII	12.88
Qy	154	VPRYDEBAHRLIMLHNLSLAKQOOLDYAIFIVQVANQTFNRKLMNVGYDVASRLPW	21.33
Ds	129	VPRAREHHRLRLLYTHPFLQROQLAAGIYVHQAGNGTFNRKLLNVGVREALRDEEN	1.88
Qy	214	QCFPIFDVLLREDDBNLYTC-PLOPRMNSVAIDKFNKYLPYSALPGISALYTDHKKI	27.72
Ds	189	DCFLFLHDVLLPBNDBNLVYCDPRGRHVAAMNFGSLSPYPOYFGVGSALYTDQYLKM	24.88
Qy	273	NGFSNDPMWGGEDDLATRTSMAGLAKTSRYPTOIARUKMIKHSTEATNPNPKCRYKMG	33.22
Ds	249	NGFPNENWGGEDDDIATRYVALAOMKLSRPTISGHIKQWYKHHGDGNGNEBNPHRPFLV	3.08
Qy	333	QYGRARTRDGLSNLKYKLVNLEKLYLTYRAVVDL	3.66
Ds	309	RTQNSWTODGANSLSYVOLLARELGLLYNITADII	34.2

RESULT 4
AEB95868
ID AEB95868 standard; protein; 393 AA

OS	Homo sapiens.
XX	
PN	US2005181437-A1.
XX	
PD	18-AUG-2005.

PF	13-APR-2005;	2005US-00105796
XX		
PR	17-JUL-1998;	98US-00118464
PR	24-APR-2002;	2002US-00132652

PA (GLYC-) GLYCOZYM APS

PI Clausen H, Bennett EP,
XX
DR WPI; 2005-563212/57.

PT New nucleic acid encoding UDP-galactose: beta-N-acetylglucosamine beta-1,4-galactosyltransferase, useful for identifying DNA polymorphism in patients.

PS Disclosure; SEQ ID NO 5; 33pp; English.

CC The present invention relates to UDP-galactose: beta-N-acetylglucosamine
CC beta-1,4-galactosyltransferase (betagal-transferases) polypeptides and
CC their encoding polynucleotides. The invention is useful for identifying
CC DNA polymorphism in patients. The present sequence is the human betagal-1-
CC T3 protein.

Query Match	35.0%;	Score 710;	DB 9;	Length 393;
Best Local Similarity	48.5%;	Pred. No. 1.2e-66;		
Matches 133; Conservative	43;	Mismatches 96;	Indels 2;	Gaps 2;

QY 95 PDNQTQQFNCQTPPHLVGPRIIV-FLDEPDEFKLEKIYPDTAAGHGMPKDCCVARRHRAII 153

DB 69 PARGGLPYDPERSPLLVGPSVSSEPPVSLAEIVERNPRVEDGGRRYPAGECEPRSRPAII 128

QY 90 RPLFPDNGTLOFCNQTPRHVGPVRFLDER-DFKTELEKIPDTHAGHGMPKDCVARN 148
 Db 80 GPRNLPDSAPSGLCDPSPFLVGLPVRKVERSGQVNLSEAVSTNPEVRBGRFAPKCKALQ 139
 QY 149 RVAIIIVPYRDRBAHLRIMLHNLHSLAKQOLDYALFIVEQVANQTFNRGKLMNVGYDAS 208
 Db 140 KVAIIIPFRNBEHLKMYLVMHPILOQOLDYGVVAINQDDEBFNRKALNVGFTEAL 199
 QY 209 RLYPMQCFIFHDVDLLPEDDRNLVTCPIQPRHMSVAIDKFNKLPYSALFGGISALTCDH 268
 Db 200 KEYDYDCVFSVDVLLPMDDBRNTYKCYSPRHLVSMDKFGRLPYNQYFGVSALSKEQ 259
 QY 269 LKINGFNSDFWGWGEGDDDLATRTSMAGLKYRSRYPTQIARYKMTKXSTEAATNPVKCRY 328
 Db 260 FTKINGFPNNYMGWGEDDDIYNRLVFKMGISRPDAVIGCKRMIRHSRDRKNEPERF 319
 QY 329 KIMGQTKRMTDGLSNLKYKLVNLEKPLYTRAVVDL 366
 Db 330 DRIATRTMTSSDGLNSLSYEVLRDTRPFLYTRITVDI 357

RESULT 7
 ADL27179
 ID ADL27179 standard; protein; 362 AA.
 AC ADL27179;
 XX
 XX
 DT 20-MAY-2004 (first entry)
 XX
 XX
 DE Chicken beta1,4-galactosyltransferase.
 XX
 KM chicken; galactosyltransferase; beta1,4-galactosyltransferase; cancer;
 KM adenocarcinoma; leukemia; lymphoma; melanoma; developmental disorder;
 KM renal tubular acidosis; anaemia; Cushing's syndrome; Wilms' tumour;
 KM aniridia; mental retardation syndrome; congenital glaucoma; cataract;
 KM reproductive disorder; infertility; endometriosis;
 KM polycystic ovary syndrome; uterine fibroid; Peyronie's disease;
 KM benign prostatic hyperplasia; autoimmune disorder; inflammatory disorder;
 KM AIDS; Addison's disease; asthma; atherosclerosis; bronchitis;
 KM Crohn's disease; atopic dermatitis; Grave's disease; multiple sclerosis;
 KM myasthenia gravis; Werner syndrome.
 KM
 OS Gallus gallus.
 XX
 XX
 PN US2003203396-A1.
 PD 30-OCT-2003.
 XX
 PF 10-JUN-2003; 2003US-00459311.
 XX
 PR 03-APR-1998; 98US-00055097.
 PR 12-AUG-1999; 99US-00373902.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Hillman JL, Guegler KJ, Corley NC, Shah P, Patterson C;
 XX WPI; 2003-900634/82.
 DR
 XX
 PT Novel human galactosyltransferase useful for treating or preventing
 CC developmental disorders such as anemia, mental retardation, cancer such
 CC as leukemia, lymphoma, reproductive disorders, autoimmune disorders.
 CC
 PS Disclosure; SEQ ID NO 7; 49pp; English.
 XX
 XX The invention relates to an isolated human galactosyltransferase
 CC polypeptide (I). (I) is useful for screening a compound for effectiveness
 CC as an agonist or antagonist of (I). (I) is useful for screening a
 CC compound that specifically binds to (I). (I) is useful for screening for
 CC a compound that modulates the activity of (I). (I) is useful for
 CC preparing polyclonal antibody (V). (V) is useful for a diagnostic test
 CC for a condition or disease associated with the expression of human
 CC galactosyltransferase (HUGA) in a biological sample. (V) is useful for

CC detecting (I) in sample and purifying (I) from a sample. (I) or its
 CC fragment or derivative is useful for treating cancer such as
 CC adenocarcinoma, leukaemia, lymphoma, melanoma, developmental disorders
 CC such as renal tubular acidosis, anaemia, Cushing's syndrome, Wilms'
 CC tumour, aniridia, mental retardation syndrome, congenital glaucoma,
 CC cataract, reproductive disorders such as infertility, endometriosis,
 CC disruptions of estrous cycle, polycystic ovary syndrome, uterine
 CC fibroids, Peyronie's disease, benign prostatic hyperplasia and
 CC autoimmune/inflammatory disorders such as AIDS, Addison's disease,
 CC asthma, atherosclerosis, bronchitis, Crohn's disease, atopic dermatitis,
 CC Grave's disease, multiple sclerosis, myasthenia gravis, Werner syndrome.
 CC The present sequence represents the amino acid sequence of chicken
 CC beta1,4-galactosyltransferase.
 CC
 SQ Sequence 362 AA;
 QY Query Match 33.5%; Score 680.5; DB 7; Length 362;
 Db Best Local Similarity 43.9%; Pred. No. 1.5e-63;
 Matches 122; Conservative 58; Mismatches 97; Indels 1; Gaps 1;
 QY 90 RPLFPDNGTLOFCNQTPRHVGPVRFLDER-DFKTELEKIPDTHAGHGMPKDCVARN 148
 Db 80 GPRNLPDSAPSGLCDPSPFLVGLPVRKVERSGQVNLSEAVSTNPEVRBGRFAPKCKALQ 139
 QY 149 RVAIIIVPYRDRBAHLRIMLHNLHSLAKQOLDYALFIVEQVANQTFNRGKLMNVGYDAS 208
 Db 140 KVAIIIPFRNBEHLKMYLVMHPILOQOLDYGVVAINQDDEBFNRKALNVGFTEAL 199
 QY 209 RLYPMQCFIFHDVDLLPEDDRNLVTCPIQPRHMSVAIDKFNKLPYSALFGGISALTCDH 268
 Db 200 KEYDYDCVFSVDVLLPMDDBRNTYKCYSPRHLVSMDKFGRLPYNQYFGVSALSKEQ 259
 QY 269 LKINGFNSDFWGWGEGDDDLATRTSMAGLKYRSRYPTQIARYKMTKXSTEAATNPVKCRY 328
 Db 260 FTKINGFPNNYMGWGEDDDIYNRLVFKMGISRPDAVIGCKRMIRHSRDRKNEPERF 319
 QY 329 KIMGQTKRMTDGLSNLKYKLVNLEKPLYTRAVVDL 366
 Db 330 DRIATRTMTSSDGLNSLSYEVLRDTRPFLYTRITVDI 357

RESULT 8
 AEB95869
 ID AEB95869 standard; protein; 362 AA.
 XX
 XX
 AC AEB95869;
 XX
 XX
 DT 20-OCT-2005 (first entry)
 XX
 XX
 DE Chicken beta-1,4-galactosyltransferase (CKI) protein, SEQ ID NO: 6.
 XX
 XX DNA polymorphism; enzyme; beta-1,4-galactosyltransferase.
 KM
 OS Gallus gallus.
 XX
 XX
 PN US2005181437-A1.
 PD 18-AUG-2005.
 XX
 PF 13-APR-2005; 2005US-00105796.
 XX
 PR 17-JUL-1998; 98US-00118464.
 PR 24-APR-2002; 2002US-00132652.
 XX
 PA (GLYC-) GLYCOTYM APS.
 XX
 PI Clausen H, Bennett EP;
 XX WPI; 2005-563212/57.
 DR GENBANK; AAB05218.
 XX
 XX New nucleic acid encoding UDP-galactose: beta-N-acetylglucosamine beta-
 PT 1,4-galactosyltransferase, useful for identifying DNA polymorphism in

PT patients.
 XX
 PS Disclosure; SEQ ID NO 6; 33pp; English.
 XX
 CC The present invention relates to UDP-galactose: beta-N-acetylglucosamine
 CC beta-1,4-galactosyltransferase (beta4Gal-transferase) polypeptides and
 CC their encoding polynucleotides. The invention is useful for identifying
 CC DNA polymorphism in patients. The present sequence is the chicken beta-
 CC 1,4-galactosyltransferase (CKI) protein.
 XX
 SQ Sequence 362 AA;
 Query Match 33.2%; Score 673.5; DB 9; Length 362;
 Best Local Similarity 43.5%; Pred. No. 8.7e-63;
 Matches 121; Conservative 58; Mismatches 98; Indels 1; Gaps 1;
 QY 90 RPLPFDNQTLCQNCOTPHLVGPIRVFLDEP-DFKLEKITYPDTHAGHGMPKCVARH 148
 DB 80 QPNNLPDSAPSGCPRPSPFLVGFPLRVFESQPVNLEEVASTNPEVEGGRFAPKCKALQ 139
 QY 149 RVALIIPYRDREAHRLIMLHSLAKQOLDYAFIVEGVANOTENRGKLMNVGYDVAS 208
 DB 140 KVALIIPFRNREBHLKWLTYMHPILQKQOLDYGVVINDGDEFPKALNVGFTEAL 199
 QY 209 RLYPMOCFIFHDYDLFPEDDRNLVTCPIQPRHMSVAIDKENYKLPYSALFGISALTQKH 268
 DB 200 KEYDYDCVFSVDVLIIPMDRNTYKCYSPRHLSVSMKDFGFLPYNQYFGVSALSKEQ 259
 QY 269 LKKINGFSNDFWNGGEDDLARTSMAGLKVSRYPTQIARYKMKIATSTATNPVKCRY 328
 DB 260 FTKINGFPNNYMGWGEDDDIYNRLVFKMGISRPDAVICKRMIRHSRDRKNEPNERP 319
 QY 329 KIMGOTRRWTRDGLSNLKYKLVNLEKPLTYRAVVDL 366
 DB 320 DRIAHTRMTSSDGLNSLSTEVLRTRFPLTYRTIVDI 357

RESULT 9
 AAB27731
 ID AAB27731 standard; protein; 264 AA.
 AC AAB27731;
 DT 30-JAN-2001 (first entry)
 XX
 DE Sequence homologous to protein fragment encoded by gene 2.
 XX
 CC Cytolethic; immunosuppressive; nootropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 CC vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 CC cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 CC neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN MO20005201-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US006059.
 XX
 PR 12-MAR-1999; 99US-0124096P.
 PR 03-DEC-1999; 99US-0168622P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI; 2000-628182/60.
 XX
 PT Novel human secreted proteins useful for diagnosis, prevention and
 PT treatment of disorders including neurological, cell proliferative,
 PT cardiovascular, autoimmune/inflammatory disorders and microbial

PT infections.
 XX
 PS Disclosure; Page 9; 427pp; English.
 XX
 CC The invention relates to the isolation of genes AAC59157-C59205 encoding
 CC the human secreted proteins AAB27682-B27730. This sequence represents a
 CC peptide fragment homologous to the protein encoded by the gene given in
 CC the descriptor line. The sequence is a search result from a BLASTX
 CC homology search. The genes and proteins are useful for preventing
 CC ameliorating or treating medical conditions, e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections
 XX
 SQ Sequence 264 AA;
 Query Match 32.9%; Score 668.5; DB 3; Length 264;
 Best Local Similarity 45.1%; Pred. No. 1.8e-62;
 Matches 119; Conservative 55; Mismatches 89; Indels 1; Gaps 1;
 QY 103 CNOTPHLVGPIRVFLDEP-DFKLEKITYPDTHAGHGMPKCVARHVAIIPYRDREA 161
 DB 1 CPDPSPLVGPPLRVESQPVNLEEVASTNPEVEGGRFAPKCKALQKVAIIPFRNRE 60
 QY 162 HLRIMHNLHSLAKQOLDYAFIVEGVANOTENRGKLMNVGYDVASRLYPMOCFIFHDV 221
 DB 61 HLTMYLYMHPILQKQOLDYGVVINDGDEFPKALNVGFTEALKEYDYDCVFSVDV 120
 QY 222 DLLPEDDRNLVTCPIQPRHMSVAIDKENYKLPYSALFGISALTQKHLLKINGFSNDFWG 281
 DB 121 DLIPMDRNTYKCYSPRHLSVSMKDFGFLPYNQYFGVSALSKEQFTKINGFPNNYMG 180
 QY 282 WGEEDDDLATRTSMAGLKVSRYPTQIARYKMKIATSTATNPVKCRYKMGOTRRWTRD 341
 DB 181 WGEEDDDIYNRLVFKMGISRPDAVICKRMIRHSRDRKNEPNERPDRIAHTRMTSSD 240
 QY 342 GLSNLKYKLVNLEKPLTYRAVVD 365
 DB 241 GLNSLSTEVLRTRFPLTYRTIVD 264

RESULT 10
 AAR55709
 ID AAR55709 standard; protein; 767 AA.
 AC AAR55709;
 DT 25-MAR-2003 (revised)
 DT 18-NOV-1994 (first entry)
 XX
 DE Glycosyltransferase hybrid.
 XX
 CC Galactosyltransferase; sialyltransferase; hybrid protein;
 CC glycosyltransferase; glycoprotein; glycolipid; oligosaccharide; YEPGSTb;
 CC glycosylation; Saccharomyces cerevisiae.
 XX
 OS Homo sapiens.
 XX
 PN MO9412646-A1.
 XX
 PD 09-JUN-1994.
 XX
 PF 15-NOV-1993; 93WO-EP003194.
 XX

CC (1); preparing (M1) a saccharide composition having a defined sequence
CC using (1); a composition prepared by (M1); a kit comprising packaging
CC material, and polypeptide comprising (1); and linking a donor into an
CC acceptor that is attached to a blood platelet using (1). (1) Is useful
CC for synthesizing a galactose-beta (1,4)-N-acetylglucosamine group,
CC glucose-beta (1,4)-N-acetylglucosamine group, N-acetylglactosamine-beta
CC (1,4)-N-acetylglucosamine group, N-acetylglactosamine-beta (1,4)-glucose
CC group, N-acetylglucosamine-beta (1,4)-N-acetylglucosamine group, mannose-
CC beta (1,4)-N-acetylglucosamine group and galactose-beta (1,4)-N-
CC acetylglucosamine-6-SO 3 group, which involves incubating a reaction
CC mixture comprising (1), with a donor and an acceptor. (1) Is useful for
CC increasing the immunogenicity of antigen and for preparing a saccharide
CC composition having a defined sequence. (1) Is useful in the preparation
CC of vaccines. This is the amino acid sequence of mouse beta(1,4)-
CC galactosidase I.
XX

Sequence 399 AA;

Query Match 32.3%; Score 667.5; DB 9; Length 399;

Best Local Similarity 45.7%; Pred. No. 4,4e-62;

Matches 123; Conservative 52; Mismatches 93; Indels 1; Gaps 1;

QY 99 TLQPCNQTTPHLYGPIRV-FLDEBDFKTLKITYPDTHAGHGMPKDCVARRVAIIYPR 157
DB 127 SLPACPESEPLVGPMLIDPNIADVLELLAKKNPEIKTGGRYSPKDCVSPHKVAIIIPFR 186
QY 158 DREAHLRIMLHNHSLAKQQLDYAIFIVEQVANQTFNRGKLMNVGYDVASRLYPMQCFI 217
DB 187 NRQEHKLYMYLHPITIQROLDYGIYINQAGDTWFRKALNIGQELKQDYNCFV 246
QY 218 FHDVDLLPEDDRNLVTCPIQPRHNSVAIDKFNKLPYSALFGGISALTQDLKKINGFSN 277
DB 247 FSDVDLIPMDRRAVRCFSQPRHISVAMDKFGFSLPYQYFGGVSAISKQFLAINGFPN 306
QY 278 DFWGGEDEDDLATRISMAGLKVSRYPTQIARFKMIGSTEATNPVVKCRXYKIMGQTKRR 337
DB 307 NYWGWGEDEDDINRRLVHKMSISRPNAAVVGRCRMIRHSRDKKNEPNPQRFRIAHTKET 366
QY 338 WTRDGLSNLKYKLVNLELKPPLYTRAVVDL 366
DB 367 MRPDGLNSLTYKYVDVQRYPLTYQITVDI 395

Search completed: January 6, 2006, 15:19:28
Job time : 135 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:16:49 ; Search time 46 Seconds
(without alignments)
688.365 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARKSLVLCV.....VLEKDCRRLRDPFPCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/ECTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	35.3	481	2	US-09-949-016-8746
2	710	35.0	393	2	US-09-118-464-5
3	710	35.0	393	2	US-10-132-652A-5
4	704	34.7	336	2	US-09-270-767-43058
5	680.5	33.5	362	1	US-09-055-097-7
6	680.5	33.5	362	1	US-09-373-902-7
7	673.5	33.2	362	2	US-09-118-464-6
8	673.5	33.2	362	2	US-10-132-652A-6
9	668.5	32.9	767	1	US-08-446-777-8
10	666.5	32.8	767	1	US-08-446-777-6
11	666	32.8	398	1	US-08-446-777-2
12	657	32.3	372	2	US-09-118-464-2
13	657	32.3	372	2	US-09-949-016-8066
14	657	32.3	372	2	US-09-949-016-8067
15	657	32.3	372	2	US-10-132-652A-2
16	649	32.0	400	2	US-09-118-464-4
17	649	32.0	400	2	US-10-132-652A-4
18	646	31.8	342	2	US-09-118-464-3
19	646	31.8	344	1	US-10-132-652A-3
20	642.5	31.6	344	1	US-09-055-097-3
21	642.5	31.6	344	2	US-09-373-902-3
22	642.5	31.6	344	2	US-09-949-016-6412
23	642.5	31.6	344	2	US-09-991-181-236
24	642.5	31.6	344	2	US-09-990-444-236
25	642.5	31.6	344	2	US-09-997-333-236
26	642.5	31.6	344	2	US-09-992-598-236
27	642.5	31.6	356	2	US-09-949-016-7638

28	583	28.7	351	2	US-09-949-016-10246	Sequence 10246, A
29	583	28.7	388	2	US-09-949-016-6551	Sequence 6551, Ap
30	577	28.4	382	2	US-09-949-016-6550	Sequence 6550, Ap
31	576	28.4	404	2	US-09-949-016-8230	Sequence 8230, Ap
32	510.5	25.1	164	2	US-09-270-767-58392	Sequence 58392, A
33	314.5	15.5	166	2	US-09-270-767-33562	Sequence 33562, A
34	314.5	15.5	166	2	US-09-270-767-48779	Sequence 48779, A
35	272.5	13.4	327	2	US-09-247-155-94	Sequence 94, Appl
36	272.5	13.4	327	2	US-09-599-3608-12	Sequence 12, Appl
37	272.5	13.4	327	2	US-09-513-989C-12	Sequence 12, Appl
38	272.5	13.4	327	2	US-09-471-276-12	Sequence 12, Appl
39	272.5	13.4	327	2	US-10-012-231A-17	Sequence 17, Appl
40	272.5	13.4	327	2	US-10-015-389A-17	Sequence 17, Appl
41	272.5	13.4	327	2	US-09-903-190-94	Sequence 94, Appl
42	272.5	13.4	327	2	US-10-006-768A-17	Sequence 17, Appl
43	272.5	13.4	327	2	US-10-015-671A-17	Sequence 17, Appl
44	272.5	13.4	327	2	US-10-015-393A-17	Sequence 17, Appl
45	272.5	13.4	327	2	US-10-011-833A-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-949-016-8746
; Sequence 8746, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8746

LENGTH: 481

TYPE: PRT

ORGANISM: Human

US-09-949-016-8746

Query Match

Best Local Similarity

Matches 134; Conservative

43; Mismatches 95; Indels 2; Gaps 2;

Score 717; DB 2; Length 481;

Pred. No. 2.2e-72;

95 PDNQTQFCNCTPPLVGPRIYV-FLDEPDKLEKIYPTDAGHGMPKDCVARRHVAII 153

157 PAPQCLPYCPRSPFLVGPVSFSPVSLAIYERNRVRPGRGYRPGCEPRRTAII 216

154 VPIRREHRLALIMLNHLSLAKOOLDAIFIVEQVANOQTNRGKLMVGVDSRLXPW 213

217 VPIRREHRLALIMLNHLSLAKOOLDAIFIVEQVANOQTNRGKLMVGVDSRLXPW 213

214 QCFIHHVDLLPEDDRLNYTC-PIOPRMSVAIDKFNKLYPSAIFGISAALTOKHLKI 272

277 DCLFLHVDLLPEDDRLNYTC-PIOPRMSVAIDKFNKLYPSAIFGISAALTOKHLKI 272

273 NGFSNDPFGWGGEDDDLTFTSMAGLKYSRYPTQIARYKMTIKHSTTEATNPVKKCYKIMG 332

337 NGFPEYWGCGEDDDLTFTSMAGLKYSRYPTQIARYKMTIKHSTTEATNPVKKCYKIMG 332

333 QTKRWRTRDGLSNLYKYLWNLKPIYTRAVVDL 366

397 RTQNSWTDGNSLTYQLALRELGLTYNTTADI 430


```
; GENERAL INFORMATION:
; APPLICANT: Clausen, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
; FILE REFERENCE: 4305/0E521
; CURRENT APPLICATION NUMBER: US/09/118,464A
; PRIORITY FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-09-118-464-6
```

```
Query Match      33.2%; Score 673.5; DB 2; Length 362;
Best Local Similarity 43.5%; Pred. No. 1.2e-67;
Matches 121; Conservative 58; Mismatches 98; Indels 1; Gaps 1;
```

```
QY 90 RPLFPDNLQFCNQTPPHLVGPIRVFLDEP-DFKLTLEKIYDPTTHAGHGMPKCVARR 148
DB 80 QPNPLPDSASGICPPDPSPLVGLRVEFQPNVLEAVSTNPEVREGKFAFKDCKALQ 139
QY 149 RVAIIVPYRDREAHRLIMLHLSLAKQQLDYAIFVEQVANQTFNRKGLMNVGYDVAS 208
DB 140 KVAIIIPFRREHLLKWLVMPIIQROQLDGVYVINDGDGEFRPAKLAVGFTLAL 199
QY 209 RLYPMOCFIFHDVLLPEDDRNLVTCPTQPRHMSVAIDKENYKLPYSATFGISALTQKH 268
DB 200 KEYDYDCFFPSVDLLPMDDRNTYKCYSPRHLVSMDKGFRLPYVQYFGVSALSKEQ 259
QY 269 LKKINGFNSDFWGMGEGDDDLATRTSMAGLKVSRYPTQIARVMKIRHSTEATNPVKKCRY 328
DB 260 FTKINGFPNNYWGEGEDDIYNRLVFKMGISRPDAVIGKCRMTIRHSRDKNEPNERF 319
QY 329 KIMGQTKRRTDGLSNLKYKLVNLEKLPYTRAVVDL 366
DB 320 DRIAHRTMTSSDGLNSLSYEVLRTRDRFPLYTRITVDI 357
```

```
RESULT 8
US-10-132-652A-6
; Sequence 6, Application US/10132652A
; Patent No. 6916649
; GENERAL INFORMATION:
; APPLICANT: CLAUSER, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; FILE REFERENCE: 04305/100E521-US1
; CURRENT APPLICATION NUMBER: US/10/132,652A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIORITY FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-10-132-652A-6
```

```
Query Match      33.2%; Score 673.5; DB 2; Length 362;
Best Local Similarity 43.5%; Pred. No. 1.2e-67;
Matches 121; Conservative 58; Mismatches 98; Indels 1; Gaps 1;
```

```
QY 90 RPLFPDNLQFCNQTPPHLVGPIRVFLDEP-DFKLTLEKIYDPTTHAGHGMPKCVARR 148
DB 80 QPNPLPDSASGICPPDPSPLVGLRVEFQPNVLEAVSTNPEVREGKFAFKDCKALQ 139
QY 149 RVAIIVPYRDREAHRLIMLHLSLAKQQLDYAIFVEQVANQTFNRKGLMNVGYDVAS 208
```

```
DB 140 KVAIIIPFRREHLLKWLVMPIIQROQLDGVYVINDGDGEFRPAKLAVGFTLAL 199
QY 209 RLYPMOCFIFHDVLLPEDDRNLVTCPTQPRHMSVAIDKENYKLPYSATFGISALTQKH 268
DB 200 KEYDYDCFFPSVDLLPMDDRNTYKCYSPRHLVSMDKGFRLPYVQYFGVSALSKEQ 259
QY 269 LKKINGFNSDFWGMGEGDDDLATRTSMAGLKVSRYPTQIARVMKIRHSTEATNPVKKCRY 328
DB 260 FTKINGFPNNYWGEGEDDIYNRLVFKMGISRPDAVIGKCRMTIRHSRDKNEPNERF 319
QY 329 KIMGQTKRRTDGLSNLKYKLVNLEKLPYTRAVVDL 366
DB 320 DRIAHRTMTSSDGLNSLSYEVLRTRDRFPLYTRITVDI 357
```

```
RESULT 9
US-08-446-777-8
; Sequence 8, Application US/08446777
; Patent No. 5641668
; GENERAL INFORMATION:
; APPLICANT: Berger, Eric G.
; APPLICANT: Matzele, Manfred
; APPLICANT: Iwanow, Svetoslav X.
; TITLE OF INVENTION: Proteins having glycosyltransferase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: CIBA-GEIGY Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: BPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/A/BE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-777-8
```

```
Query Match      32.9%; Score 668.5; DB 1; Length 767;
Best Local Similarity 35.4%; Pred. No. 1.5e-66;
Matches 140; Conservative 74; Mismatches 132; Indels 49; Gaps 6;
```

```
QY 14 LVLCAVLLLVHAMIV-----KIPSLYENTLIGSSTLIADVDAMEAVLGNTA----- 60
DB 26 LVAVCALHLHGVTLVYVLAGRDL SRLPQ-----GVSTPLQGGSSNAAAIIGSSGLRTG 80
QY 61 -----STSDLLDITWNSTFSPISSEVNOTSFMEDIRPLLPFNQLOQC 103
DB 81 GARPPLGASQPRPGDSSPVVDGPGPASNLTSV-----PVPHTTALSLPAC 130
```


MOLECULE TYPE: protein
US-08-446-777-2

Query Match 32.8%; Score 666; DB 1; Length 398;
Best Local Similarity 35.7%; Pred. No. 1e-66;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHAMI-----KIPSYENLTIGSTLADVAMEAVLGNTA-----60
26 LVAVCALHLGVTLVYLAGDLSRLPOL-----VGVSTPLGGSSNSAAGSSGGLRIG 80
61 -----STSDLLDTWNSITFSPISEVNOTSFMEDIRPILFPDNOTLQPC 103
81 GAPPPEPLGASQPRFGDSSPVVDGSPASMLTSV-----PVPHTLALSPLAC 130
104 NQPPHNLNGIRFLEDP--DFKLEKTYPTHAGGCMPCQVARRVAIIVPYRDEAH 162
131 PESSPLVGMLEFNNPVDLELVAQNPVVKMGGRYAPEDCVSPHKVAIIFRRQEH 190
163 LRIMLNHLSLAKQOOLVAIFIVEQVANOTFNRGKLMNVGVDAASRLYPMOCFIFHDV 222
191 LKWLTYLHVLROQUDIGIYINQAGDIFNRKALINVGPOELKDIDYTCFVFSVDV 250
223 LLEPDDRNLVTCPIOPRHMSVAIDKENYKLPSAIFGIGSALTCKDLKINGFSNDPFGW 282
251 LIPMNDHNAVRCFSOPRHISVAMDKFGFSLPYVQFGVGSALSKQOFLTINGFPNNYMGW 310
283 GGGDDDLARTSMAGLKVSRYPQIARVYKMKIKSTEAITNVNCKRYKIMQTRKRTROG 342
311 GGGDDDIIFNNLVFRGMSISRPNAVGCRCMIRHSRDKKBNPNORFDRLAHTEYMLSDG 370
343 LSNLYKLVNLEKPLTYTRAVVDL 366
371 LNSLTYQVDLVQRYRPLTYQTIVDI 394

RESULT 12
US-09-118-464-2
Sequence 2, Application US/09118464A
Patent No. 6558934
GENERAL INFORMATION:
APPLICANT: Clausen, Henrik
APPLICANT: Bennett, Eric Paul
TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
TITLE OF INVENTION: Beta-1,4-Galactosyltransferase, Beta4gal-12
FILE REFERENCE: 4305/08521
CURRENT APPLICATION NUMBER: US/09/118,464A
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-09-118-464-2

Query Match 32.3%; Score 657; DB 2; Length 372;
Best Local Similarity 37.2%; Pred. No. 9.6e-66;
Matches 135; Conservative 71; Mismatches 135; Indels 22; Gaps 6;

12 KSLVLCAVLLVHAMIYKIPSYENLTIGSS--TLADVAMEAVLGNTASTSDLLDTW 70
14 KAVLLLCILHFLVAVILY--FDVYAQHLAFSPRSARGPAHALHPAASSSSSSCSRPNA 72
71 NSTFSPISEVNOTSFMEDIRPILFPDNOTLOFCNQTPPHLVGPIRVFLDEP--DFKTIEXI 129
73 TASSSGLPPEVPSA-----LPGPTAPTLPPCPDSPPGVLVGRLLIFTSFMLERVQRE 124
130 YPPTHAGGCMPCQVARRVAIIVPYRDEAHLRIMLNHLSLAKQOOLVAIFIVEOV 189
125 NPGVLMGGRYTPPDCPTPAQTVAVIIPFRHREHLRYMLHYPILRORLRYGVYVINOH 184
190 ANQTFNRGKLMNVGY-----DVASRLYPMOCFIFHDVLLPEDDRNLVTCPIOPRHMSV 243

185 GEDTFNRAKLVNMGFILEALKEDAA-----YDCFISVDVLYPMDDRNLYRCGDQPRHFAI 239
244 AIDKFNKYKLPSAIFGIGSALTCKDLKINGFSNDPFGMGGGEDDLATTSMAGLKVSRX 303
240 AIDKFGFRLLPYAGYFGVSGLSKAQFLRINGFPNEWMGGBDDIIFNLSLGMKISR 299
304 PTOIARYKMKIKSTEAITNVNCKRYKIMQTRKRTROGSLNLYKLVNLEKPLTYTRAV 363
300 DIRIGRYMKIKDRDKHNEPNPQRFKIONTKLTKMRDGDIGISVRYQVLEVSQPLFTNIT 359
364 VDL 366
360 VDI 362

RESULT 13
US-09-949-016-8066
Sequence 8066, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8066
LENGTH: 372
TYPE: PRT
ORGANISM: Human
US-09-949-016-8066

Query Match 32.3%; Score 657; DB 2; Length 372;
Best Local Similarity 37.2%; Pred. No. 9.6e-66;
Matches 135; Conservative 71; Mismatches 135; Indels 22; Gaps 6;

12 KSLVLCAVLLVHAMIYKIPSYENLTIGSS--TLADVAMEAVLGNTASTSDLLDTW 70
14 KAVLLLCILHFLVAVILY--FDVYAQHLAFSPRSARGPAHALHPAASSSSSSCSRPNA 72
71 NSTFSPISEVNOTSFMEDIRPILFPDNOTLOFCNQTPPHLVGPIRVFLDEP--DFKTIEXI 129
73 TASSSGLPPEVPSA-----LPGPTAPTLPPCPDSPPGVLVGRLLIFTSFMLERVQRE 124
130 YPPTHAGGCMPCQVARRVAIIVPYRDEAHLRIMLNHLSLAKQOOLVAIFIVEOV 189
125 NPGVLMGGRYTPPDCPTPAQTVAVIIPFRHREHLRYMLHYPILRORLRYGVYVINOH 184
190 ANQTFNRGKLMNVGY-----DVASRLYPMOCFIFHDVLLPEDDRNLVTCPIOPRHMSV 243
185 GEDTFNRAKLVNMGFILEALKEDAA-----YDCFISVDVLYPMDDRNLYRCGQPRHFAI 239
244 AIDKFNKYKLPSAIFGIGSALTCKDLKINGFSNDPFGMGGBDDDLATTSMAGLKVSRX 303
240 AIDKFGFRLLPYAGYFGVSGLSKAQFLRINGFPNEWMGGBDDIIFNLSLGMKISR 299
304 PTOIARYKMKIKSTEAITNVNCKRYKIMQTRKRTROGSLNLYKLVNLEKPLTYTRAV 363
300 DIRIGRYMKIKDRDKHNEPNPQRFKIONTKLTKMRDGDIGISVRYQVLEVSQPLFTNIT 359
364 VDL 366
360 VDI 362

```

RESULT 14
US-09-949-016-8067
; Sequence 8067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8067
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8067

Query Match      32.3%; Score 657; DB 2; Length 372;
Best Local Similarity 37.2%; Pred. No. 9.6e-66;
Matches 135; Conservative 71; Mismatches 135; Indels 22; Gaps 6;

QY 12 KSLVLCVALLVHAMIKIPSLYENLTIGSS-TLIAVDAMEAVLGNASTSDLDLTW 70
DB 14 KAVLLLCILHFLVAVILY-FDYVAQHIAFFSFRSARPAHALHPAASSSSSSNGSRPA 72
QY 71 NSTSPPISEVNQTSFMEIDIRPLFPDNOQLQFCNQTPPHVLGPIVFLDER-DFKTLKXI 129
DB 73 TASSSGLPEVPSA-----LPGPTAPTLPPCPDPSPPGLVGRLLIEFTSPMPRLRVORE 124
QY 130 YPDTHAGGHPKDCVARRHVAIVPYRDREAHILMNLHSLAKQQLDVAIFIVEGV 189
DB 125 NGVIMGGRYTPDCTPAQVAVILPFRRHREHLRYWLHYLHPILRQRLRGVAVINOH 184
QY 190 ANQTFNRKGLMNVGY-----DVASRLYPWOCFIFHDVDLLPEDDRNLTYTCPIQPRHMSV 243
DB 185 GEDTFNRKALNVGFLKEDAA-----YDCFIFSDVDLVPMDDRLNLYRCDDQPRHFAI 239
QY 244 AIDKFNKYLPSAIFGIGSALTQDHLKKINGFSNDFWGWGEGEDDLATRTSMAGLKVSRY 303
DB 240 AMDKGFRLPYAGYGVGSGLSKAQFLRINGFPNEYWGWGEGEDDLIFNRIISLTGMKISR 299
QY 304 PTQIARYMKIKHSTETATNPVNCRCRYKINGQTRRTROGLSNLYKLVNLELKLPLYTRAV 363
DB 300 DIRIGRYMKIKDRKNEPNPQRTKIQNTLTKMRDGISVRYQVLEVSROPPLFTNIT 359
QY 364 VDL 366
DB 360 VDI 362

RESULT 15
US-10-132-652A-2
; Sequence 2, Application US/10132652A
; Patent No. 6916649
; GENERAL INFORMATION:
; APPLICANT: CLAUSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; FILE REFERENCE: 04305/100521-US1
; CURRENT APPLICATION NUMBER: US/10/132,652A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIOR FILING DATE: 1998-07-17

```

```

; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-652A-2

Query Match      32.3%; Score 657; DB 2; Length 372;
Best Local Similarity 37.2%; Pred. No. 9.6e-66;
Matches 135; Conservative 71; Mismatches 135; Indels 22; Gaps 6;

QY 12 KSLVLCVALLVHAMIKIPSLYENLTIGSS-TLIAVDAMEAVLGNASTSDLDLTW 70
DB 14 KAVLLLCILHFLVAVILY-FDYVAQHIAFFSFRSARPAHALHPAASSSSSSNGSRPA 72
QY 71 NSTSPPISEVNQTSFMEIDIRPLFPDNOQLQFCNQTPPHVLGPIVFLDER-DFKTLKXI 129
DB 73 TASSSGLPEVPSA-----LPGPTAPTLPPCPDPSPPGLVGRLLIEFTSPMPRLRVORE 124
QY 130 YPDTHAGGHPKDCVARRHVAIVPYRDREAHILMNLHSLAKQQLDVAIFIVEGV 189
DB 125 NGVIMGGRYTPDCTPAQVAVILPFRRHREHLRYWLHYLHPILRQRLRGVAVINOH 184
QY 190 ANQTFNRKGLMNVGY-----DVASRLYPWOCFIFHDVDLLPEDDRNLTYTCPIQPRHMSV 243
DB 185 GEDTFNRKALNVGFLKEDAA-----YDCFIFSDVDLVPMDDRLNLYRCDDQPRHFAI 239
QY 244 AIDKFNKYLPSAIFGIGSALTQDHLKKINGFSNDFWGWGEGEDDLATRTSMAGLKVSRY 303
DB 240 AMDKGFRLPYAGYGVGSGLSKAQFLRINGFPNEYWGWGEGEDDLIFNRIISLTGMKISR 299
QY 304 PTQIARYMKIKHSTETATNPVNCRCRYKINGQTRRTROGLSNLYKLVNLELKLPLYTRAV 363
DB 300 DIRIGRYMKIKDRKNEPNPQRTKIQNTLTKMRDGISVRYQVLEVSROPPLFTNIT 359
QY 364 VDL 366
DB 360 VDI 362

```

Search completed: January 6, 2006, 15:24:10
Job time : 48 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 8, 2006, 22:26:55 ; Search time 279 Seconds
(without alignments)
3010.334 Million cell updates/sec

Title: US-10-661-430-2
Perfect score: 1152
Sequence: 1 atgagcttcgcattcgtgc.....acttccaacgtgttttag 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4637633 seqs, 364532575 residues

Total number of hits satisfying chosen parameters: 9275266

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgnt2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgnt2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgnt2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgnt2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgnt2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgnt2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgnt2_6/prodata/1/pubpna/US11_NEW_PUB.seq2:*
8: /cgnt2_6/prodata/1/pubpna/US11_NEW_PUB.seq3:*
9: /cgnt2_6/prodata/1/pubpna/US11_NEW_PUB.seq4:*
10: /cgnt2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	155.4	13.5	1964	7	US-11-102-240-41
2	150	13.0	4646	6	US-10-955-054A-120
3	138.8	12.0	951	7	US-11-136-527-1055
4	138.8	12.0	951	7	US-11-136-527-5151
5	53.2	4.6	16822	6	US-10-995-561-13225
6	53.2	4.6	18705	6	US-10-995-561-13401
7	50.8	4.4	2507	6	US-10-750-185-30398
8	50.8	4.4	2507	6	US-10-750-185-30398
9	50.2	4.4	201	6	US-10-995-561-22169
10	50.2	4.4	201	6	US-10-995-561-68468
11	41.8	3.6	1728	6	US-10-750-623-43601
12	41.8	3.6	1728	6	US-10-750-623-43601
13	34.6	3.0	3458	6	US-10-750-185-34973
14	34.6	3.0	3458	6	US-10-750-623-34973
15	33.6	2.9	879	6	US-10-467-657-4833
16	32.6	2.8	14656	7	US-11-121-086-68
17	32.6	2.8	190882	7	US-11-121-086-69
18	32.2	2.8	1954	6	US-10-750-185-55399
19	32.2	2.8	1954	6	US-10-750-623-55399
20	31.8	2.8	1559	6	US-10-750-185-32663
21	31.8	2.8	1559	6	US-10-750-623-32663
22	31.8	2.8	2704	6	US-10-750-185-57973
23	31.8	2.8	2704	6	US-10-750-623-57973

C 24	31.4	2.7	1375	6	US-10-750-185-58979	Sequence 58979, A
C 25	31.4	2.7	1375	6	US-10-750-623-58979	Sequence 58979, A
C 26	31.4	2.7	3721	6	US-10-750-185-25033	Sequence 25033, A
C 27	31.4	2.7	3721	6	US-10-750-623-25033	Sequence 25033, A
C 28	31.2	2.7	1380	6	US-10-750-185-45020	Sequence 45020, A
C 29	31.2	2.7	1380	6	US-10-750-623-45020	Sequence 45020, A
C 30	31.2	2.7	1655	6	US-10-750-185-44286	Sequence 44286, A
C 31	31.2	2.7	1655	6	US-10-750-623-44286	Sequence 44286, A
C 32	31.2	2.7	151152	6	US-10-750-185-45728	Sequence 45728, A
C 33	31.2	2.7	1453	6	US-10-750-623-45728	Sequence 45728, A
C 34	31	2.7	1453	6	US-10-750-623-45728	Sequence 45728, A
C 35	30.8	2.7	1027	7	US-11-136-527-2988	Sequence 1068, Ap
C 36	30.8	2.7	1113	6	US-10-793-626-1063	Sequence 1068, Ap
C 37	30.8	2.7	1147	7	US-11-136-527-877	Sequence 877, Ap
C 38	30.8	2.7	1147	7	US-11-136-527-4973	Sequence 4973, Ap
C 39	30.8	2.7	2986	6	US-10-793-626-3517	Sequence 3517, Ap
C 40	30.8	2.7	5925	7	US-11-000-688-1358	Sequence 1358, Ap
C 41	30.8	2.7	154548	7	US-11-121-086-33	Sequence 33, Appl
C 42	30.6	2.7	79134	7	US-11-124-368A-4924	Sequence 2924, Ap
C 43	30.4	2.6	201	7	US-11-124-368A-4901	Sequence 4901, Ap
C 44	30.4	2.6	201	7	US-11-124-368A-12257	Sequence 12257, A
C 45	30.4	2.6	600	6	US-10-750-185-21776	Sequence 21776, A

ALIGNMENTS

RESULT 1
US-11-102-240-41
Sequence 41, Application US/11102240
GENERAL INFORMATION:
Publication No. US20050260647A1
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
FILE REFERENCE: P3230R1C106C
CURRENT APPLICATION NUMBER: US/11/102,240
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: 10/063662
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 199-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 41
LENGTH: 1964
TYPE: DNA
ORGANISM: Homo Sapien
US-11-102-240-41

Query Match 13.5%; Score 155.4; DB 7; Length 1964;
Best Local Similarity 51.8%; Pred. No. 3.3e-42;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY	421	CCTAAGATTGTTGTTGCAAGCATGTTGTTATTTATGTCCTATAGAGATGCTGAA	480
DB	461	CCTCAGAGATGTAAGCTTTACAGAGGTCGTCCTCTCCACCGGAAGAGAG	520
QY	481	GCACATTGAGAAATATGCTCCACATTTTGCACTGTTGCTGCCAACAACAAATTGAC	540
DB	521	AAACACCTGATATACCTGCTGGAACATGCTATCCCTTCACAGGACGAGCTGAT	580
QY	541	TATGCAATTTTCAATGTTGAGCAATGCGAATCAGCTTAATCCGGGAACTAATG	600
DB	581	TATGCAATTTTCAATGTTGAGCAATGCGAATCAGCTTAATCCGGGAACTAATG	640

```
QY 601 AACGTTGATACGACGATGATCAGCTCTACCCATGAGGAGCTTTCATCTTTCATGAT 660
DB 641 AATGTGGCTATATAGAGCCCTCAAGAGAAATTTGGACCTGCTTATATTCACAGAT 700
QY 661 GTTCGATTTACGTCGGAGAGATGACCGTAACTGTACACGTGTCAATTCAACCACTCAT 720
DB 701 GTGGAACCTGTACCCGAGATGACTTTAACCTTTACAGATGTGAGAGACATCCCAAGCAT 760
QY 721 ATGAGTGTACCGATGATATTAATTCATTAATTAACCTTCATATTCGGCATCTTCGGCGA 780
DB 761 CTGTGTGTGTGGAGAACACACACTGGGTACAGTTACGTTACAGTGATATTTTGGGGGT 820
QY 781 ATCAGTGACATTAACAAAGATCACTGAAGAAATCAATGATTTTCATGATATTTTGG 840
DB 821 GTTACTGCTCTTAAGCAGAGAGAGTTTTCAGAGTGATGATTTCTTCAACACTACTG 880
QY 841 GGTGGGGCGGAGAGACGACGATTTGGCGACAGAAATCATGATGGCTGACTGAAAGTT 900
DB 881 GGATGGGGAGGCGAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAAATGAATTT 940
QY 901 TCAAGATATCCGACACAAATTTGACAGATATTAATGATTAAGCATTCGACGAGAGCAG 960
DB 941 TCCGGGCCCCCTGCTGAAGTGGTAAATATACATGCTTTCACACTAGACCAAGGC 1000
QY 961 AATCCAGTTATATATGCGCTCAAAATATATGAGGCCAAAGCGCCGATGACACGT 1020
DB 1001 AATGAGGTGAACCGACGAGAGGATGAGCTCTTACCAAGTGTACGAGTCTGGAAGCA 1060
QY 1021 GACGCGCTTAAGCATCTGAAGTATTAAGCTGTAAATGGAATTTGAAGCTCTTACACT 1080
DB 1061 GATGGGTGTAGTGTGTTCTTATTAATATAGTATCTGTGAGACAAATCCTTTATATATC 1120
QY 1081 CGAGCCGTCTCGATTT 1097
DB 1121 AACATCACAGTGATTT 1137
```

RESULT 2

```
US-10-955-054A-120
; Sequence 120, Application US/10955054A
; Publication No. US2005026420A1
; GENERAL INFORMATION:
; APPLICANT: PUSZTAI, LAJOS
; APPLICANT: SYMMANS, W. FRASER
; APPLICANT: HESS, KENNETH R.
; APPLICANT: AYERS, MARK
; APPLICANT: STEC, JAMES
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
; FILE REFERENCE: UTXC:880US
; CURRENT APPLICATION NUMBER: US/10/955,054A
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 120
; LENGTH: 4646
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-955-054A-120
```

Query Match 13.0%; Score 150; DB 6; Length 4646;
Best Local Similarity 56.8%; Pred. No. 4,1e-40;
Matches 276; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

```
QY 405 CGGTGACATGGAATGCTTAAGATTGTGTGCAAGGATCGTTCCTATATTTGCGC 464
DB 559 CGGAGGTGACATGGAAGCTTCTGATTCATGCTCGGTGGAAGGTGCGATCTTATCCC 618
QY 465 CTATAGAGATCGTGAAGACATTTGAGAAATAGCTCCACAATTTGACGTGTGCTGCG 524
DB 619 CTTCGGAAACCGCACAGACCTCCAGTCTCTTTAGACACCTGCTTCCATGCTTCA 678
QY 525 CAAACACAAATTGACTATGCAATTTTCATTTGAGACAAATGCGAATCGATCGATTAA 584
```

```
DB 679 GCGCAGCGCTTCAGATTGCAATTTTATGTGTGAGAACAGTTGATACCAACCCCTTAA 738
QY 585 TCGCGGAAACATTAAGACGTGATAGACGATAGCATCAGCCTCTACCATGCGAGTG 644
DB 739 TGAAGCATGCTTTTCAAGGTGGCTTTTCAAGAGCAATGAAGACTTGGATTGGACTG 798
QY 645 CTTCATCTTTCATGATGTGATTTTACCTGCGCAGAGATGACCGTAACTGTACAGCTGTC 704
DB 799 TTTGATTTTTCATGATGTGATACATACCGAAGATGTCCCACTATTTATGATGTGG 858
QY 705 AATTCAACCAAGTCATATGATGTAGCATTCATTAATTTCAATTTAACTTCATATTC 764
DB 859 ACAGATGCCGAGGCAATTTTGCAACCAAAATGATATGATATGTATCTGCTTCTATAC 918
QY 765 GCGATCTTCGGCGGAATGATGATGACCTAACAAGATCACTGAAGAAATCAATGATTT 824
DB 919 CGAGTTCCTTTGGCGAGTATGATGTGCTTAAACATGGAACAATTTGGAAATCAATGCTT 978
QY 825 TTGCAATGATTTTGTGGGTGGCGGAGAGACGACGATTTTGGCGACGAGAACATCAT 884
DB 979 TCTATATGCTTCTTGCGGTTGGGGGTGAGAAATGACGACTCTGGAACAGATACAGAA 1038
QY 885 GCGCTG 890
DB 1039 TGCAGG 1044
```

RESULT 3

```
US-11-136-527-1055
; Sequence 1055, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentn version 3.2
; SEQ ID NO 1055
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-11-136-527-1055
```

Query Match 12.0%; Score 138.8; DB 7; Length 951;
Best Local Similarity 48.5%; Pred. No. 9,8e-37;
Matches 390; Conservative 7; Mismatches 404; Indels 3; Gaps 1;

```
QY 294 GACTCTTCAATTTGTATATCAGACACTCCCACTCTGTCGACCCATCCGTATTTCT 353
DB 132 GACTCTGCTGCTTGTGCTCGAGAGATCCCGCTGCTGTGTGTCATGTGATGACTT 191
QY 354 CGATGAGCC---CGACTTCAAAACTCTCGAGAAATCTATCCGAGACGACGACGCGATG 410
DB 192 TAATATTTCTGTGGAATCTGAGACTTTTGGCAAGAAACCCAKAATAAAGATGGCGG 251
QY 411 ACATGGAATGCTTAAGATTTGTTCGAAGCATGCTGTGTCTATTTATGTGCTTATAG 470
DB 252 SCGTACTTCCCAAGAGMGTATCTCCCTCACAAGGTGGCCATATATCCATTTCCG 311
QY 471 AGATGTGAAGCATTTGAGAAATATGCTCCACAATTTGCACTGCTGTGCTGCCAAACA 530
DB 312 TAACCGGAGGAGACCTCAAAATACGTGCTGTATTTTTCATCCAGCTCTTCAGCGCA 371
QY 531 ACAATTTGACTATGCAATTTTCATTTGAGAGCAAGTGGCAATCGACGTTTATGCGCG 590
DB 372 GCAACTGATGATGCAATCTCAATCAATTCAGGCTGAGAGACCAATGTTTAAATGAGC 431
QY 591 GAAACTAATGAAGTTGATACGATGACGATGACGCTCTACCATGAGCATGCTTCAT 650
```

```

Db 432 TAAGTGTGCTCAAGCTTTGGCTTTCAGAGAGCCCTTGAAGACTATGACTACAACTGGCTTGT 491
Qy 651 CTTTCATGATGTGCTGATTTACTGCCCCGAGATGACCGTAACTGTACAGCTGTCCAAATTC 710
Db 492 GTTCAGTATGTGAGACCTCATTTCCATGTAGTACCATATATGCKKACAGGTGCTTTTCA 551
Qy 711 ACCAGTCAATATGAGTGTGAGATGATGATTAATTCATTTAATTCATTCATTTCCGCGCAT 770
Db 552 GCGACGGCATATTTCTGTGCAATGAGCAAGTTGGGTTTACCTGCTTACGTTCAATTA 611
Qy 771 CTTCCGCGGAATCACTGACATCAAAAGATCACTGAGAAATCAATGATGATTTTCGA 830
Db 612 TTTTGAGGTGTCTCGTCTCAGTAAACAGAGTTCTTACCATCAATGATTTTCCTTA 671
Qy 831 TGAATTTTGGGGTGGGGCGGAGAGACGACATTTTGGCAGAGAACATCGATGGCTGG 890
Db 672 TAATTACTGGGGCTGGGGAGAGAGATGATGACATTTTAAACAGATTGATTCATTAAGG 731
Qy 891 ACTGAAAGTTTCAAGTATCCGACACAAATTCGAGATTAATAATGATTAAGCATCTCGAC 950
Db 732 CATGCTATATATACGCCCAATGCTGTGTAGGCRGTGTGCAATCATCGGACTCAAG 791
Qy 951 GGAAGCGACGAATCCAGTTAATAATGCGCTACAAATATATGAGCCAAACGAGCGCG 1010
Db 792 AGACAAGTTTATATGAGCCCAACCTCAGAGTTTGAACCGGATCGCACATCAAGAAAC 851
Qy 1011 ATGACACGTGACGGCTTAAGCAATCTGAATTAAGCTCTGTAATCTGAAATTTGAAGCC 1070
Db 852 GATGGCGCTTGAATGTTTGAATCTTACTACAGAGGTGTGACATACAGATGATACC 911
Qy 1071 TCTCTACACTGAGCGCTGTGCGA 1094
Db 912 GTTATATACCAAAATCACAGTGA 935

```

RESULT 4

```

US-11-136-527-5151
; Sequence 5151, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5151
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-5151

```

```

Query Match 12.0%; Score 138.8; DB 7; Length 951;
Best Local Similarity 48.5%; Pred. No. 9, 8e-37;
Matches 390; Conservative 7; Mismatches 404; Indels 3; Gaps 1;

```

```

Qy 294 GACTCTTCAATTTCTGTAATCAAGACACTTCCCACTGTGCGAGCCATCCGTATTTCT 353
Db 132 GACTCTGCTGCTTGTGCTTCCGAGAGATCCCGCTGCTGTGGTCCCATGTGATGACTT 191
Qy 354 CGATGAGCC---CGACTTCAAACTCTCGAGAAATCTATCCGAGACACGACGCGGTG 410
Db 192 TAAATATCTGTGATCTGAGCTTTTGGCAAGAAAGAACCAKAGATTAAGATGGCGG 251
Qy 411 ACATGGAATGCTTAAGATTTGTGTCAAGGACATCGTGTCTATTATTGTGCTTATAG 470
Db 252 SGTGTACTTCCCAAGMCTGTATCTCCCTCAAGAGTGGCCATCATTAATCCATTCG 311

```

```

Qy 471 AGATCGTAAACACATTTGAGATATATGCTCCACATTTGCACTGCTGCGCAACA 530
Db 312 TAACCGGAGAGACACCTCAAAATCTGGCTGATATTTGATTCAGTCTTCAAGCGCA 371
Qy 531 ACAATTTGACATATGCAATTTTCAATTTGTGAGCAAGTGGCGAATCAGAGTTTATTCGCG 590
Db 372 GCAATCTGATGTGACATCTACATCATCAATCAGGCTGAGACACATGTTTAAATCGAGC 431
Qy 591 GAAATTAATGAACCTTTGATACAGATGACATACAGCCTTACATCCATGGAGCTGTAT 650
Db 432 TAAGTGTCTCAACCTTTGCTTCAAGAGCGCTTGAAGACTATGACTCACTGCTTGT 491
Qy 491 CTTTCATGATGTGATTTACTGCCCCGAGATGACCGTAACTGTACAGTGTCCAAATTC 710
Db 432 GTTCAGTATGTGACCTCATTTCCATATGATGACCATATATGCKKACAGTGTCTTCA 551
Qy 711 ACCAGTCAATATGAGTGTGAGGATGATTAATTCATTAATTAATTCATTAATTCGCGCAT 770
Db 552 GCGACGCAATATTTCTGTGCAATGAGCAAGTTGGGTTTACCTGCTTACGTTCAATTA 611
Qy 771 CTTCCGCGGAATCACTGACATCAAAAGATCACTGAGAAATCAATGATGATTTTCGA 830
Db 612 TTTTGAGGTGTCTCGTCTCAGTAAACAGATTCCTTACCATCAATGATTTTCCTTA 671
Qy 831 TGAATTTTGGGGTGGGGCGGAGAGACGACATTTTGGCGACGAGAACATCGATGGCTGG 890
Db 672 TAATTACTGGGGCTGGGAGAGAGATGATGACATTTTAAACAGATTAAGTTCAATAAG 731
Qy 891 ACTGAAAGTTTCAAGTATCCGACACAAATTCGAGATTAATAATGATTAAGCATCTGAC 950
Db 732 CATGCTATATATACGCCCAATGCTGTGTAGGCRGTGTGCAATCATCGGACTCAAG 791
Qy 951 GGAAGCGACGAATCCAGTTAATAATGCGCTACAAATATATGAGCCAAACGAGCGCG 1010
Db 792 AGACAAGTTTATATGAGCCCAACCTCAGAGTTTGAACCGGATCGCACATCAAGAAAC 851
Qy 1011 ATGACACGTGACGGCTTAAGCAATCTGAATTAAGCTCTGTAATCTGAAATTTGAAGCC 1070
Db 852 GATGGCGCTTGAATGTTTGAATCTTACTACAGAGGTGTGACATACAGATGATACC 911
Qy 1071 TCTCTACACTGAGCGCTGTGCGA 1094
Db 912 GTTATATACCAAAATCACAGTGA 935

```

RESULT 5

```

US-10-995-561-13225/c
; Sequence 13225, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: C1001559
; CURRENT FILING DATE: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13225
; LENGTH: 16822
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(16822)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-2)
US-10-995-561-13225

```

```

Query Match 4.6%; Score 53.2; DB 6; Length 16822;
Best Local Similarity 63.1%; Pred. No. 1, 1e-06;
Matches 82; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

QY	750	TAAACCTTCATATTCGGCGATCTTCGGGGAAATCAGGCACTTAAACAAAGATCAGCTGA	809
Db	13331	TAGCCTCCCGACCCCGAGTACTTCGGAGGAGTCTCAGCACTTATCTTCGACCGATACCT	132727
QY	810	GAATATCATGATTTTTCGATGATTTTGGGGTGGGGCGAGAGGATCGAGATTGGC	869
Db	13271	GAATATGAATGGCTTCCCAATGATATCTGGGGCTGGGGTGGTAGGATGACGACATTGC	13212
QY	870	GACGAGAAC	879
Db	13211	TACCAAGTCA	13202

```

RESULT 6
US-10-995-561-13401
: Sequence 13401, Application US/109395561
: Publication No. US20050272054A1
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: CL001559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FASTSeq For Windows Version 4.0
: SEQ ID NO 13401

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(18705)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
; IS-10-995-561-13401

```

Query Match	4.6%	Score 53.2	DB 6	Length 18705
Best Local Similarity	63.1%	Pred. No. 1.2e-06		
Matches 82; Conservative		0; Mismatches 48;	Indels 0;	Gaps 0;

OY	750	TAAACCTTCATATTGCGGCATCTCGGGCGGAATAGAGCACTAACAAAGATCACCCTGAA	809
Dd	10281	TAGCCTCCCGTACCCCCAGTACTTTGGAGGAGTGCTGACGACTTAATCTGTGACCATGTAACCT	10340
OY	810	GAAATCATATGATTTTTGCAATGATTTTTGGGGTGGGGCGGAGAGGACAAGATTTGGC	869
Dd	10341	GAAAGTAGATGGCTTCCCACATGAACTCTGGGGCTGGGGGTGGTGAAGATGACCACTTGC	10400

```

QY      870 GACGAGACA 879
          |||
Db      10401 TACCAGGTCA 10410

```

```

RESULT 7
US-10-750-185-30398
; Sequence 30398, Application US/10750185
; Publication NO. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFEID, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750.185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 649222

```

```

; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 30398
;
; LENGTH: 2507
;
; TYPE: DNA
; ORGANISM: Bovine
;
US-10-750-185-30398

```

Query Match	4.4%	Score 50.8	DB 6	Length 2507
Best Local Similarity	62.7%	Pred. No. 2e-06		
Matches 79	Conservative 0	Mismatches 47	Indels 0	Gaps 0

Oy 754 CTTCCAAATTTGGGGGAACTTGGGGGAAATCAGTGCATCAACAAAGATCACCCTGAAGAA 813
 Db 1007 CTCCTGATATCCCAAGTACTTTGGGGGGGTCTCAAGCGCTCACTCTGCACAGTACTCTGAAG 1066
 Oy 814 ATCAATGATTTTGGATGATATTTTGGGGTGGGGCCGAGAGACAGACGATTTTGGCCAG 873
 Db 1067 ATGAATGGCTTCCCAATGAAATATCTGGGGCTGGGGGTGTGAGATGACGACATTTGTACC 1128

QY	874	AGAACA	879
Db	1127	AGGTCA	1132

RESULT 8
US-10-750-623-30398
; Sequence 30398, Application US/10750623
; Publication No. US20050287531A1

```

1  APPLICANT: MMI GENOMICS, INC.
2  APPLICANT: DENISE, Sue K.
3  APPLICANT: KERR, Richard
4  APPLICANT: ROSENFELD, David
5  APPLICANT: HOLM, Tom
6  APPLICANT: BATES, Stephen
7  APPLICANT: FANTIN, Dennis
8  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
9  FILE REFERENCE: MM1100-1
10 CURRENT APPLICATION NUMBER: US/10/750,623
11 CURRENT FILING DATE: 2003-12-31
12 PRIOR APPLICATION NUMBER: US 60/437,482
13 PRIOR FILING DATE: 2002-12-31
14 NUMBER OF SEQ. ID NOS: 64922
15 SOFTWARE: PatentIn version 3.1
16 SEQ. ID NO 30398
17
18 LENGTH: 2507
19
20 TYPE: DNA
21 ORGANISM: Bovine
22 US-10-750-623-30398

```

Query Match	4.4%	Score 50.8	DB 6	Length 2507
Best Local Similarity	62.7%	Pred. No. 2.2e-06		
Matches 79; Conservative	0	Mismatches 47	Indels 0	Gaps 0

QY	754	CTTCCATATTTGGGCGATCTTTCGGCGGATTCAGTGCATTAACAAAGATCACCCTGAAGAAA	813
Db	1007	CTCCCGTATCCCAAGTACTTTGGCGGGGCTTCAGCGCTCCTCCTCGACAGCAGTACTCTGAAG	1066
QY	814	ATCATGATATTTGGAATGATTTTGGGGGTTGGGCGGAGAGACACGATTTGGCGAG	873
Db	1067	ATGATAGGCTTCCCAATGATCTGGGCGTGGGGTGGTGAAGGATGACGACTTGTCTAC	1126
QY	874	AGACACA	879
Db	1127	AGGTCA	1132

RESULT 9
US-10-995-561-22169/c
; Sequence 22169, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```

1  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
2  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
3  TITLE OF INVENTION: DETECTION AND USES THEREOF
4  FILE REFERENCE: CL001559
5  CURRENT APPLICATION NUMBER: US/10/995,561
6  CURRENT FILING DATE: 2004-11-24
7  NUMBER OF SEQ. ID NOS: 85702
8  SOFTWARE: FASTSEQ for Windows Version 4.0
9  SEQ ID NO 22169
10 LENGTH: 201
11 TYPE: DNA
12 ORGANISM: Homo sapiens
13 US-10-995-561-22169

```

Query Match	4.4%;	Score 50.2;	DB 6;	Length 201;
Best local Similarity	57.5%;	Pred. No. 6	8e-07	
Matches	88;	Conservative	1;	Mismatches 64; Indels 0; Gaps 0

QY	550	TTCAATTGTGGAGCAAGTGGCGAATCAGACGTTTAATCGCGGAAACTAATGAACGTTGA	609
Db	154	TCATTATCTTCAGAGCTGGAATGGAACATTAAACAGGGCAAACTGTGTAAATGTTGGG	95
QY	610	TACGAGTGTAGCATCAGCGCTCTAACCCATGGCAGTGTCTCATCTTTCATGATGTCGATT	669
Db	94	GTCGCGAGAGGCCCTGTGGTATGAAGAATGGGAGCTCGCTTCTTTCGACGATGTGAACCTC	35
QY	670	CTGCCGAAGATGACCGCTAACCTGTACACGTT	702
Db	34	TTGCCAGAAAATATACCAACATCTGTATGTGTGT	2

```

RESULT 10
US-10-995-561-68468
; Sequence 68468, Application US/10995561
; Publication No. US20050272054A1
;
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
;
NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68468
;
LENGTH: 201
;
TYPE: DNA
;
ORGANISM: Homo sapiens
;
US-10-995-561-68468

```

Query Match	4.4%;	Score 50.2;	DB 6;	Length 201;
Best Local Similarity	57.5%;	Pred. No. 6.8e-07;		
Matches	88;	Conservative 1;	Mismatches 64;	Indels 0; Gaps 0
QY	550	TTGATTGTGGAGCAGTCGCGAATCAGACGTTTAAATCCGCGGAAATCATGACGTTGGA	609	
Db	48	TCATTATCTTCCAGCGCTGGAATGGAACATTTAACAGGCAAACTGTAAAGTTGGG	107	
QY	610	TACGACGTACATCAGCCTCTAACCCATGCGAGTCTTCATCTTCATGATGTCGATTTA	668	
Db	108	GTGCGAGAGGCCCTGTGCTGATGATGAGAGTGGGACTGCTGTTCTTGACGATGTGGA	167	
QY	670	CTGCGGAGATGACCGCTAACCGTATGACAGT	702	
Db	168	TTGCCAGAAATATACACATCTGTATGTGTGT	200	

```

RESULT 11
US-10-750-185-43601
; Sequence 43601, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:

```

```

1  APPLICANT: MMI GENOMICS, INC.
2  APPLICANT: DENISE, Sue K.
3  APPLICANT: KERR, Richard
4  APPLICANT: ROSENFELD, David
5  APPLICANT: HOLM, Tom
6  APPLICANT: BATES, Stephen
7  APPLICANT: FATTIN, Dennis
8  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
9
10 FILE REFERENCE: MM1100-2
11
12 CURRENT APPLICATION NUMBER: US/10/750,185
13
14 CURRENT FILING DATE: 2003-12-31
15
16 PRIOR APPLICATION NUMBER: US 60/437,482
17
18 PRIOR FILING DATE: 2002-12-31
19
20 NUMBER OF SEQ ID NOS: 64922
21
22 SOFTWARE: PatentIn version 3.1
23
24 SEQ ID NO 43601
25
26 LENGTH: 1728
27
28 TYPE: DNA
29
30 ORGANISM: Bovine 19866880501107
31
32 US-10-750-185-43601

```

	Query Match	Best Local Match	Similarity	3.6% Pred. No. 0.002	Score 41.8	DB 6	Length 1728	Matches 88	Conservative 0	Mismatches 77	Indels 0	Gaps 0
QY	699	GTCGCAATTCACACGATGATAGAGTGAAGCATGATTAATTCATTATTAACCTCC						758				
DB	1165	GAGAACAAAGAGACTTCGTCTAACCTTCCTGCTCTCTCTCTTCAGGTACG						1224				
QY	759	ATATTCGGCCATCTTCGGCGGAATCACTGCCTTAACAAAATATCACTCGAAGAAATCA						818				
DB	1225	TTAAAGTGATATTTTGGGGGTGTACTGTCCCTTAATGCGAAGACAATTTTCAAGTGAA						1284				
QY	819	TGATTTTCGATGATTTTGGGGGTGGGGCGAGAGAGACAGCA						863				
DB	1285	TGATTTCTTAACAACCTACTGGGGATGGGGAGCGGAAGACGATGA						1329				

```

RESULT 12
US-10-750-623--43601
Sequence 43601, Application US/10750623
Publication No. US20050287531A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43601
LENGTH: 1728
TYPE: DNA
ORGANISM: Bovine 19866880501107
US-10-750-623--43601

```

Query Match	3.6%;	Score 41.8;	DB 6;	Length 1728;
Best Local Similarity	53.3%;	Pred. No. 0.002;		
Matches 88;	Conservative 0;	Mismatches 77;	Indels 0;	Gaps 0;

Oy	699	GTGTCGAATTCACACCGATCATAGACATGACGATGATTAATTCATTAATCAATCC	758
Db	1165	GAGACAAAAGACATTGCTGTACCTCTTCCCTCTCTTCCTCTTCAGAGTTGCG	1222
Oy	759	ATAATTCGGCATCTTCGGCGGATCATGCACTAACAAGATCATCCTGAAGAAATCA	818

Db 1225 TTAGAGTGAATATTTGGGGGTGTACTGCGCTAAGTCGAGACAAATTTTCAAGTGAA 1284
Qy 819 TGGATTTTGAATGATTTTGGGGGTGTGGGCGGAGAGACGACGA 863
Db 1285 TGGATTTCTTAACAACTACTGGGATGGGAGGCGAAGACGATGA 1329

RESULT 13

US-10-750-185-34973
; Sequence 34973, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34973
; LENGTH: 3458
; TYPE: DNA
; ORGANISM: Bovine 19866881077545
US-10-750-185-34973

Query Match 3.0%; Score 34.6; DB 6; Length 3458;
Best Local Similarity 67.1%; Pred. No. 0.91;
Matches 49; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 630 CTACCATGCGAGTCTTCATCTTTCATGATGCGATTTCTGCGCCGAGATGACCGTAA 689
Db 3359 CTCACCTTAACAACCTCTTCATCTTCAGTACGTGACCTGCTCCCATGATGACCGCA 3418
Qy 690 CCTGTACACGTGT 702
Db 3419 CCTGTACCGCTGT 3431

RESULT 14

US-10-750-623-34973
; Sequence 34973, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34973
; LENGTH: 3458
; TYPE: DNA
; ORGANISM: Bovine 19866881077545
US-10-750-623-34973

Query Match 3.0%; Score 34.6; DB 6; Length 3458;
Best Local Similarity 67.1%; Pred. No. 0.91;
Matches 49; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 630 CTACCATGCGAGTCTTCATCTTTCATGATGCGATTTCTGCGCCGAGATGACCGTAA 689
Db 3359 CTCACCTTAACAACCTCTTCATCTTCAGTACGTGACCTGCTCCCATGATGACCGCA 3418
Qy 690 CCTGTACACGTGT 702
Db 3419 CCTGTACCGCTGT 3431

RESULT 15

US-10-467-657-4833
; Sequence 4833, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNAN Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 4833
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4833

Query Match 2.9%; Score 33.6; DB 6; Length 879;
Best Local Similarity 57.7%; Pred. No. 0.83;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 787 GCACCTAACAAAGATCACTGTAAGAAATCAATGATTTTCCAAATGATTTTGGGGTGG 846
Db 406 GCACCTTTGAAGGAAGCGCCGGAATAATTCATTGTGTTGGAGCATTTGTCAATTGAA 465
Qy 847 GCGGAGAGACGACGATTTGGGACGAGACATGATGCGCTGG 890
Db 466 AGCGGCGACGAAACCTATAGGCAATTGAAAACGCGATTGGATGG 509

Search completed: January 9, 2006, 01:00:42
Job time : 282 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 8, 2006, 22:31:00 ; Search time 746 Seconds
(without alignments)
3421.685 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSLLVLCV.....VDLEKDCRRRLRDRPFCF 383

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ .p2n.model -DEV=xlh
-DB=N_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10661430 @GCN 1.1 727 @runat_06012006_123134_6226 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*
1: geneeqn1980s.*
2: geneeqn1990s.*
3: geneeqn2000s.*
4: geneeqn2001as.*
5: geneeqn2001bs.*
6: geneeqn2002as.*
7: geneeqn2002bs.*
8: geneeqn2003as.*
9: geneeqn2003bs.*
10: geneeqn2003cs.*
11: geneeqn2003ds.*
12: geneeqn2004as.*
13: geneeqn2004bs.*
14: geneeqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	1152	12	ADM32969 Nucleotid
2	718	35.4	2002	3	AACT7836 Human can
3	717	35.3	1938	2	AA242120 Human end
4	717	35.3	1949	13	ACN37523 Tumour-as

ALIGNMENTS

RESULT 1		5	709	34.9	1839	4	ABLI4209	Abli4209 Drosophil
ADM32969	ID	6	684	33.7	1853	2	AAV73011	Aav73011 Human adu
XX	ADM32969 standard; cDNA; 1152 BP.	7	684	33.7	1853	6	ABO92036	Abog92036 Human pol
XX		8	675	33.2	4093	4	ABLI4208	Abli4208 Drosophil
AC	ADM32969;	9	668.5	32.9	2304	2	AAQ66892	Aaq66892 Glycosylt
XX		10	668	32.9	2615	2	AAQ03464	Aaq03464 cDNA enco
DT	17-JUN-2004 (first entry)	11	667.5	32.9	3923	2	AAQ03463	Aaq03463 cDNA enco
XX		12	667.5	32.9	3950	13	ADT99580	Adt99580 Mouse bet
DB	Nucleotide sequence of a beta1,4-N-acetylglactosaminyltransferase.	13	667.5	32.9	3953	1	ACN45213	Acn45213 Mouse MRN
XX		14	666	32.8	1158	3	AAA53309	Aaa53309 Beta 1,4
XX		15	666	32.8	1158	6	ABK90350	Abk90350 cDNA enco
XX		16	666	32.8	1192	2	AAK80151	Aak80151 Human bet
XX		17	666	32.8	1194	11	ADZ47902	Adz47902 Galactose
XX		18	666	32.8	1197	10	ADP68084	Adp68084 Human bet
XX		19	666	32.8	1197	10	ADP18447	Adp18447 Human bet
XX		20	666	32.8	1197	12	ADQ78357	Adq78357 Human bet
XX		21	666	32.8	1265	2	AAQ31434	Aaq31434 Encodes a
XX		22	666	32.8	4072	12	ADQ87369	Adq87369 Human tum
XX		23	666	32.8	4080	11	ADZ47903	Adz47903 Galactose
XX		24	666	32.8	4162	10	ADJ56362	Adj56362 Human CDN
XX		25	665.5	32.8	2304	2	AAQ66891	Aaq66891 Glycosylt
XX		26	665	32.7	1265	2	AAQ66889	Aaq66889 Galactose
XX		27	663.5	32.7	1116	13	ADQ91373	Adq91373 ManII-Gal
XX		28	661	32.5	1119	10	ADJ94853	Adj94853 Novel NOV
XX		29	661	32.5	1500	10	ADP18449	Adp18449 N-acetyl19
XX		30	660.5	32.5	1191	14	ADZ69854	Adz69854 Human Gal
XX		31	658	32.4	1116	14	ABE95896	AbE95896 Human bet
XX		32	658	32.4	1119	10	ADP73104	Adp73104 Ebgr-1,4
XX		33	658	32.4	1890	13	ADR20717	Adr20717 Human bet
XX		34	658	32.4	1890	13	ACN37366	Acn37366 Tumour-as
XX		35	658	32.4	2027	9	ACA63280	AcA63280 DNA encod
XX		36	658	32.4	2027	10	AAD59475	Aad59475 Human bet
XX		37	658	32.4	2027	14	ABE95864	AbE95864 Human bet
XX		38	657.5	32.4	3194	12	ADQ24917	Adq24917 Human sof
XX		39	656.5	32.3	905	13	ADQ80003	Adq80003 Bovine be
XX		40	656.5	32.3	2163	12	ADQ78359	Adq78359 alpha-1,3
XX		41	655.5	32.3	1068	10	ADP68134	Adp68134 Human gal
XX		42	655.5	32.3	1152	10	ADP68140	Adp68140 Human gal
XX		43	655.5	32.3	1152	10	ADP68086	Adp68086 Human gal
XX		44	655.5	32.3	6406	14	ADZ69841	Adz69841 Plasmid p
XX		45	655	32.3	1197	13	ADQ79988	Adq79988 Human bet

PR 13-SEP-2002; 2002US-0411242P.
XX
XX (CDMM/) CUMMINGS R D.
PA (KAWA/) KAWAR Z.
PI Cummings RD, Kawar Z;
XX
XX WPI: 2004-270055/25.
DR P-PSDB; ADM32968.
XX
XX
XX New purified approximatelyb4 acetyl[galactosamyl] transferase that is
PT substantially free of other proteins, useful for transforming or
PT transfecting host cells for producing substantially pure forms of the
PT enzyme.
XX
XX
XX Claim 5; Fig 1; 51pp; English.
XX
XX The present sequence encodes a beta1,4-N-acetyl[galactosamyl]transferase,
CC designated beta4GALNACT. The enzyme is required for the biosynthesis of
CC animal cell glycoproteins. The enzyme functions to synthesise the
CC lactinac or LDN motif GalNAcbeta4GlcNAc-R. This motif is glycoprotein
CC hormones produced by the pituitary gland, and is a major marker of
CC glycoconjugates made by parasitic and non-parasitic invertebrates and may
CC be implicated in host immune responses to infection. The protein and
CC polynucleotides are useful for transforming or transfecting host cells
CC for producing substantially pure forms of the enzyme, or in vitro, for
CC formation of a LDN structure on proteins or peptides.
XX
XX Sequence 1152 BP; 314 A; 275 C; 259 G; 304 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 3.53e-223 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-661-430-1 (1-383) x ADM32969 (1-1152)
QY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuLysSerLeuLeuValLeuCyAlaVal 20
DB 1 ATGGCTTTTGGTCAATTTGGCAGTCGCCAGACTCAAGTCGTTGCTCGTACTTTGGCCGTT 60
QY 21 LeuLeuLeuValHisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThrIle 40
DB 61 CTCTATTAGTTCAAGCAATGATTTATTAAGATTCATCGCTTTCAGAACTTACTATC 120
QY 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla 60
DB 121 GGCCTCCCGACCTTATTTGCGAGCTCGACCAATGAGGAGAGTCTCGGGAATACGGCT 180
QY 61 SerThrSerAspAspLeuLeuAspThrTTrpAsnSerThrPheSerProIleSerGluVal 80
DB 181 TCCACTTCGGATGATCTACTTGAATACGTGGAATTCACGTTTCCACCGAATTTCTGAAGTT 240
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
DB 241 AATGAGACTGTTTATGAGAGACATTCGTCATTCCTGTTCCCGCAACACGACTCTT 300
QY 101 GlnPheCysAsnGlnThrProHisLeuValGlyProIleArgValPheLeuAspGlu 120
DB 301 CAATTCCTGTAATCGACACCTCCCACTGCTGCGACCAATCGGTATTCCTCGAGAG 360
QY 121 ProAspPheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyGlnHisGlyMet 140
DB 361 CCGGACTTCAAAATCTCGAGAAATCTATCCGAGACACGCGGTGGACATGAGATG 420
QY 141 ProLysAspCysValAlaArgHisArgValAlaIleLeuValProTyrAspAspArgGlu 160
DB 421 CTTAAGGATTTGTGTGAGGAGCATCGTGTGCTATTTATGTCCTTAATGAGATCGTGA 480
QY 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAsp 180

DB 481 GCACATTTGAGATATATGCTCCACATTTGACATCGTTGCTCGCAACCAACATTTGAC 540
QY 181 TyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMet 200
DB 541 TATGCAATTTTATTTGAGAGCAAGTGGGAAATCAAGACTTTAAATCGCGGAAACTAAATG 600
QY 201 AsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAsp 220
DB 601 AACGTTTGATACGACGTACATCACCCCTTACCATGCGCATGCTTCACTTCAATGAT 660
QY 221 ValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHis 240
DB 661 GTCGATTTTACTCGCCGAAGATGACCGTAACCGTACACGTGTCATTCATACACGTCAT 720
QY 241 MetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyGly 260
DB 721 ATGAGGTGAGCCATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
QY 261 IleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTyr 280
DB 781 ATCAGTGCATTAACAAAGATCACTCGAGAAATCAATGATTTTCCGATGATTTTTCG 840
QY 281 GlyTyrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysVal 300
DB 841 GGTGGGGGCGAGAGGACGACGATTTGGCGACGAGAACTCATGCTGAGCTGAGTGAAGTT 900
QY 301 SerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThr 320
DB 901 TCAGATATTCGACACAAATTTGACGATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
QY 321 AsnProValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTyrThrArg 340
DB 961 AATCCGATTAATTAATGCGCGCTACAAATTAATGCGCAACGAGCGCGAGTGAACAGT 1020
QY 341 AspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThr 360
DB 1021 GACGGCCTTAAGCAATCTGAATTAAGTCTGTAATCTGGAATTTGAAGCTCTCTCACT 1080
QY 361 ArgAlaValValAspLeuLeuGluLysAspCysArgArgGluLeuArgArgAspPhePro 380
DB 1081 CGAGCGGTGTCGATTTGCTCGAAAAGACTGCGCGGAGCTGCGAAAGGACTTTCCA 1140
QY 381 ThrCysPhe 383
DB 1141 ACCTGTTTT 1149

RESULT 2
AAC77836
ID AAC77836 standard; cDNA; 2002 BP.
XX
AC AAC77836;
XX
XX 08-FEB-2001 (first entry)
DE
XX Human cancer associated gene sequence SEQ ID NO:230.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
KW antidiabetic; antihistaminic; antineumatic; antitubercular; antiviral;
KW antihistaminic; antihistaminic; antihistaminic; antihistaminic;
KW dermatological; neuroproliferative; thrombolytic; coagulant; nociceptive;
KW vasorelaxant; antihistaminic; antihistaminic; antihistaminic;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
XX
PN W0200055350-AL.
XX
XX 21-SEP-2000.

XX 08-MAR-2000; 2000MO-US005882.
XX
XX 12-MAR-1999; 99US-012427DP.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI, 2000-587533/55.
XX P-PSDB; AAB43627.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX
XX Claim 1; Page 801-802; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnary; immunomodulator;
XX antidiabetic; antiaesthetic; antineumatic; antibacterial; antiviral;
XX antiinflammatory; antitumor; antiallergic; antiparasitic; antipain;
XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
XX neurotropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilization of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancer, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 2002 BP; 428 A; 576 C; 527 G; 469 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 9,32e-72 Length: 2002
Score: 718.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.35% Indels: 2
DB: 3 Gaps: 2

US-10-661-430-1 (1-383) x AAC7836 (1-2002)

QY 95 Proaapraenglthleuglnphecyabanglthrpriorhisleuvalaileu 114
DB 470 CCARCTCCCTCAAGGCTCTGCTCCAGAACATCTCTCTAGTGGTCTGTG 529
QY 115 Argval---PheleuaspriubProaaprelysthrleuglnylettrProasphr 133
DB 530 TCGGTCCTTTAGCCCACTGCTGCAAGATTGTGAGGGAATCCCGGGTA 589
QY 134 HIsalaglyglnHIsiglymerProlysaerCyvalaAlaargHIsarGValaIailelle 153
DB 590 GAACAGAGGGGGCGGATCCGCTGAGGTGTGAGCCCGCTCCGAAACACCATCATT 649
QY 154 ValProtyrArgaAspArgGluAlaHIsleuAglIlemerleuHIsaenleuHIsereleu 173
DB 650 GTGCTCACTGTCGCGGAGGACCACTGCGCTGCTCTACCACTGACCCCTTC 709
QY 174 leuAlaIyGlnGlnleuaspTyraAlaIepheleuAlaIeGlnValaAlaenglnthr 193
DB 710 TTGCAAGCCGACAGCTTCTTATGCACTTATGTCATCACCACAGGCTGAAATGGAACA 769
QY 194 PheaasrAglylyleuHIsaenValaIeGlytyrAspValaIIsarArgleuTyrrProtrp 213

DB 770 TTTAACAAGGGCAAACTGTTGAACGTTGGGCTGGAGAGCCCTGCGTATGAAGATGG 829
QY 214 GlnCysephelelepheneIsaerValaIspleuProGuaaPaAspArgAsnleuTythr 233
DB 830 GACTGCTGTTCTTGCACAGATGGAACCTTTCGAGAAATGACCAAACTGTATGTG 889
QY 234 Cys---ProIleGlnProArgHIsaerValaAlaIeAspLyPheAsnTyrrysleu 252
DB 890 TGTGACCCCGGGGAGACCCGCGCATGTCGCTGTATGAACAAGTTTGGATACAGCCTC 949
QY 253 ProTyseraIaIlephelygIyIIsereAlaIeuthrlysaerHIsleuIysIle 272
DB 950 CCGTACCCCGCACTTCTGAGAGTCTCAGCAGCTTACTCTCTGACCAAGTACCTGAAAGTG 1009
QY 273 AenglyPheAspAaapPheTTPGlyTyrrpIyGlyIuAspAAspLeuAlaIthrArg 292
DB 1010 AATGCTTCCCAATGAATACTGGGGCTGGGTGTGAGATGACGACATTTGCTACAGG 1069
QY 293 ThserMeAlaIyLeuIyValaIeAspArgTyrrProthrGlnIleAlaArgTyrrysmet 312
DB 1070 GTGGCCCTGGGTGGATGAATCTCTGCGCCCGCCATCTGTAGACACATATAGATG 1129
QY 313 IleyshIsaerThrguAlaIthrAspProValaenlyCyAsgTyrrysIleMetGly 332
DB 1130 GTGAAGCACCGAGAGATTAAGGCAATGAGGAAATCCCGACAGTTTGAACCTCTGCTC 1189
QY 333 GlnThrlysaAspArgTyrrThrArgAspGlyleuAsnleuTyrrlyleuValaen 352
DB 1190 CGTACCCGAAATTCCTGAGCAGAGATGGATGAATCTGACATGACATGAGTGTGCT 1249
QY 353 leuglnleuTyrrProleuTyrrThrArgAlaValaIeAspLeu 366
DB 1250 CGAGAGCTGGGGCTCTTTATACCAACTCACAGAGACATT 1291
RESULT 3
AAZ42120
ID AAZ42120 standard; cDNA, 1938 BP.
XX
XX AAZ42120;
XX
XX 31-JAN-2000 (first entry)
XX
XX Human endometrium tumour cDNA derived EST 140.
XX
XX DE
XX XX Endometrium; human; tumour; cancer; anticancer; cytostatic;
XX KM EST; treatment; uterine; gene therapy; expressed sequence tag; ss.
XX
XX OS Homo sapiens.
XX
XX PN DE19817948-A1.
XX
XX 21-OCT-1999.
XX
XX PF 17-APR-1998; 98DE-01017948.
XX
XX PR 17-APR-1998; 98DE-01017948.
XX
XX (META-) METAGEN GBS GENOMFORSCHUNG MBH.
XX
XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E,
XX DR WPI, 1999-591957/51.
XX
XX PT New nucleic acid sequences expressed in uterine cancer tissues, and
XX PT derived polypeptides, for treatment of uterine and endometrial cancer and
XX identification of therapeutic agents.
XX
XX PS Claim 3; Page 273; 444pp; German.
XX
XX CC This invention describes novel human nucleic acid (cDNA) sequences (A),
XX CC that are highly expressed in uterine tumour tissue and which have
XX CC anticancer and cytostatic activity. (A) are used (i) for recombinant
XX CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)

Alignment Scores:

Pred. No.: 1.17e-71 Length: 1949
 Score: 717.00 Matches: 134
 Percent Similarity: 64.60% Conservative: 43
 Best Local Similarity: 48.91% Mismatches: 95
 Query Match: 35.30% Indels: 2
 DB: 13 Gaps: 2

US-10-661-430-1 (1-383) x ACN37523 (1-1949)

```

OY 95 ProApsanGlnThrLeuGlnPheCySaengInThrProProHisLeuValGlyProIle 114
DB 467 CCAGCTCTCCCAAGGCTGCTCCCTACTGTCAGAACATCTCTCTGTAGTGGCTGTG 526
OY 115 ArgVal----PheLeuAspGluProAspPheLysThrLeuGlnLysIleLeuProAspThr 133
DB 527 TCGGTCCTCTTACGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 586
OY 134 HisAlaGlyGlyIleGlyMetProLysAspCySaValAlaArgHisArgValAlaIleIle 153
DB 587 GAACCAAGGGGGCCGGTACCCGCTGAGGTTGTGAGCCCGCTCCGGAACAGCCATTCAT 646
OY 154 ValProTyrrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
DB 647 GTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
OY 174 LeuAlaLysGlnGlnLeuAspTyrrAlaIlePheIleValGlnGlnValAlaAsnGlnThr 193
DB 707 TTGCAGCCGACAGCAGCTGCTGCTTATGCACTATGATCATCCACAGCTGGAATGAGACA 766
OY 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrrAspValAlaSerArgLeuTyrrProIle 213
DB 767 TTTTACAGGGGCAAACTGTGTGAACCTGTGGGCGAGAGCCCTGCTGATGAAGAGTGG 826
OY 214 GlnCyPheIlePheHisAspValAspLeuProGluAspAspArgAsnLeuTyrrThr 233
DB 827 GACTCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 886
OY 234 Cys---ProIleGlnProAspHisMetSerValAlaIleAspLysPheAsnTyrrLysLeu 252
DB 887 TGTGACCCCGGGGACCCCGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946
OY 253 ProTyrrSerAlaIlePheGlyIleSerAlaLeuThrLysAspHisLeuLysIle 272
DB 947 CCGTACCCCACTACTGCTGAGAGAGTCTAGACATTACTCCGACCAAGTACTGAAAGATG 1006
OY 273 AsnGlyPheSerAspAspPheTyrrGlyIleGlyAspAspAspLeuAlaThrArg 292
DB 1007 AATGCTCTCCCAATGAATTAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
OY 293 ThrSerMetAlaGlyLeuLysValSerArgTyrrProThrGlnIleAlaArgTyrrLysMet 312
DB 1067 GTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
OY 313 IleLysHisSerThrGlyLysAsnProValAsnLysCySaArgTyrrLysIleMetGly 332
DB 1127 GTGAAGACCCGAGAGATTAAGGCAATGAGAAATCCCAACATTTACCTCTGCTGCTG 1186
OY 333 GlnThrLysAspArgTyrrThrArgAspGlyLeuSerAsnLeuLysTyrrLysLeuValAsn 352
DB 1187 CGTACCCCAATTCCTGAGCGCAAGATGGGATGAATCACTACATACAGATTGCTGCTGCT 1246
OY 353 LeuGlnLeuLysProLeuTyrrThrArgAlaValValAspLeu 366
DB 1247 CGAGAGCTGGGGCTCTTTATACCAACATCACAGACGACATT 1288

```

RESULT 5

ABL14209
 ID ABL14209 standard; cDNA; 1839 BP.

AC ABL14209;
 XX
 DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37109.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ss.
 XX OS Drosophila melanogaster.

W0200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US009231.

23-MAR-2000; 2000US-0191637P.

11-JUL-2000; 2000US-00614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI. 2001-656860/75.

P-PSDB; ABB70106.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPI at ftp.wpi.int/pub/published_pcr_sequences

Alignment Scores:

Pred. No.: 8.92e-71 Length: 1839
 Score: 709.00 Matches: 150
 Percent Similarity: 58.03% Conservative: 56
 Best Local Similarity: 42.25% Mismatches: 129
 Query Match: 34.91% Indels: 20
 DB: 4 Gaps: 7

US-10-661-430-1 (1-383) x ABL14209 (1-1839)

```

OY 28 IleTyrrLysIleProSerLeuTyrrGluAsnLeuThr-----IleGlySerSerThr 44
DB 228 GTGACCAAGTATGCTCATATCTACGGGAACGCTGACAGATGAGCGGAGGAGTGAA 297
OY 45 LeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAsp 64
DB 298 --GCATCCAGCTGCTCCCGCTTCCCGCTTATCAAAAAGACAGAGCGGAGCAG 354
OY 65 AspLeuLeuAspThrThrPheSerThrPhe---SerProIleSerGluValAsnGlnThr 83
DB 355 GAGCTCAATGGCGGAGCCCACTTACCACTTAAGAACTGTGATTTGCCAGGCAATCTTACT 414
OY 84 SerPheMetGluAspIleArgProIleLeu-----PheProAspAsn 97
DB 415 TCCATTCCACAAGACTTAACCGCTTCTGCTGCGGACAAAGAAATTTTGGCCCGGCA 474
OY 98 Gln-----ThrLeuGlnPheCySaengInThrProProHisLeuValGly 112
DB 475 CAGAAATCCACATCCGCTCTTGCACAACTGATCCGATCCCGATGAGTGGGA 534

```

```

QY 113 Pro11eArgValPheLeuAerGluProAerPheLeuGlu--Lys11eTyPro 131
DB 535 CCATACACGCCCAACAGACACGAGTCACTGAGCTTATTGAGCGGAGCTTGAGCCT 594
QY 132 AepThiHia1aGly1yH1sGlyMetProLysAerCyVal1AlaArgH1sArgVal1Ala 151
DB 595 CTTTTCGCCCTGAGCGCTTGCAGCTGAAACATGCAATGCCAGCATCAGTGGCT 654
QY 152 Ile11eValProTyArgAerAerGlu1AlaH1sLeuArg11eMetLeuH1sAerLeuH1s 171
DB 655 ATGTGTGCTCCCTCGGATGATACGCCATCTATTACTTTCCGCGCAATCTCAC 714
QY 172 SerLeuLeu1Ala1yG1ng1nLeuAerTyra11ePheH1eValG1ng1nVal1AlaAer 191
DB 715 CCAATTCTGTAGAGAGCGGCATCGCCTATCGCATTTTCAATTGAGAGCAACCAACGG 774
QY 192 G1nThrPheAerArgGlyLysMetAerVal1GlyTyArgVal1AlaSerArgLeuTy 211
DB 775 AAGCCCTTAAATCGGCTGCATGATGAACATTGTTAGAGCGCTTAAAGCTGTAC 834
QY 212 ProTyG1nCyPheH1ePheH1sAerVal1AerLeuProG1nAerAerAerLeu 231
DB 835 CAGTGGGATGTTTATATTCACGATGTCATCTTCTTGGACGACCGCAATCTT 894
QY 232 TyTrThrCyPhePro11eG1nProAerG1H1sMetSerVal1Ala1eAerLysPheAerTyLys 251
DB 895 TACAACCTGTACAGTCAAGCGCGACACATGTCAGTGTAGACACCGCAATCTTACAG 954
QY 252 LeuProTySerAla1ePheGly1y1eSerAlaLeuThryAerH1sLeuLys 271
DB 955 TTCCCTTATCGATCAATATTTGGAGGTTCGCAATGACGCGTACAGCTTCAAGGCC 1014
QY 272 ILAaerGlyPheSerAerAerPheTrpGlyTrpGlyG1nAerAerAerAerAerAer 291
DB 1015 G1TAaerG1TTCGAACTGCTTTCGCGGCGCGGACGAGTGAAGCAATGTCAC 1074
QY 292 ArgThrSerMetAlaGlyLeuLysVal1SerArgTyProThrG1nLeu1AlaArgTyLys 311
DB 1075 AGGTTGAAGACGCAAACTTATTCATTCAGATTCGCGTCAACATAGCTCGTACAG 1134
QY 312 Met11eLysH1sSerThrGlu1AlaThrAerProVal1AerLysCyVal1y1eMet 331
DB 1135 ATGCTGAAGCATCAGAAAGGCAATCTTAAAG-----CCCTTGAAGAACTTA 1185
QY 332 GlyG1nThryAerArgTyTrpThrArgAerGlyLeuSerAerLeuLysTyLysLeuVal 351
DB 1186 CAGAAATGCAATGACCAAAATAGAAAGAGATGAAATCACTGATAAAGTATCTATCTAC 1245
QY 352 AserLeuGluLeuLysProLeuTyTrThrArgAlaVal1AerLeu 366
DB 1246 AGCATCAAAACAATTCCAACTTTCAGTGTACTTACAGAGCTA 1290
RESULT 6
AAV73011
ID AAV73011 standard; CDNA; 1853 BP.
XX
AC AAV73011;
XX
DT 15-MAR-1999 (first entry)
XX
DE Human adult brain secreted protein ges1_1 cDNA.
XX
KM Secreted protein; protein factor; human; ges1_1; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 154..975
FT /tag= a
FT /note= "polynucleotides comprising nucleotides 154-972
FT and 1-341 are also claimed"
XX
PN MO9842741-A2.

```

```

XX
PD 01-OCT-1998.
XX
PF 25-MAR-1998; 98WO-US005972.
XX
PR 25-MAR-1997; 97US-00825145.
PR 24-MAR-1998; 98US-00046881.
XX
PA (GEMV ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MO;
DR P-PSDB; AAW82006.
XX
PT New polynucleotides encoding secreted human proteins - are derived from
PT human foetal brain, adult testes, adult brain, or placenta cDNA
PT libraries, potentially useful as, e.g. immunostimulants or
PT antiinflammatory agents.
XX
PS Claim 35a; Page 85-86; 113pp; English.
XX
CC This cDNA sequence codes for novel human secreted protein ges1_1 (see
CC AAW82006). The cDNA was isolated from a human adult brain cDNA library
CC using methods which are selective for cDNAs encoding secreted proteins,
CC or was identified as encoding a secreted or transmembrane protein on the
CC basis of computer analysis of the amino acid sequence of the encoded
CC protein. The nucleotide sequence shows homology to some known database
CC sequences. ges1_1 is a full-length clone and can be obtained from the
CC deposit clone ATCC 98371 as a 1850 bp EcoRI/NcoI fragment. The invention
CC provides polynucleotides (see AAV73000-12), derived from human foetal
CC brain, adult testis, adult brain and adult placenta cDNA libraries, that
CC encode novel secreted proteins (see AAW81997-07). These can be isolated
CC from composite clone ATCC 98371 using specific probes (see AAV73013-22).
CC The polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals (no supporting data
CC are given). Suggested activities include nutritional, immune stimulating
CC (e.g. as vaccines) or suppressing, haematopoiesis regulating, tissue
CC growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour invasion
CC suppressor, and tumour inhibition activities. The polynucleotides are
CC also stated to be useful for gene therapy
XX
SQ Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;
XX
Alignment Scores:
Pred. No.: 6,79e-68 Length: 1853
Score: 684.00 Matches: 133
Percent Similarity: 63.87% Conservative: 42
Best Local Similarity: 48.54% Mismatches: 97
Query Match: 33.68% Indels: 3
DB: 2 Gaps: 2
US-10-661-430-1 (1-383) x AAV73011 (1-1853)
QY 95 ProAerAerG1nThryLeuG1nPheCyAerG1nThrProProH1sLeuVal1GlyPro11e 114
DB 358 CCAAGCTCTCAAGAGTCTGCCCTACTGTCCAGAAAGATCTCTCTTGAAGGGGTCCGTG 417
QY 115 ArgVal---PheLeuAerGluProAerPheTyTrThrLeuG1nLys11eTyProAerThr 133
DB 418 TCGGTGCTTTCAGCCAGTGCATCATCTGCGAGAGATTCGTGAGGAAATCCCGGGA 477
QY 134 H1s1aGly1yH1sGlyMetProLysAerCyVal1AlaArgH1sArgVal1Ala1e11e 153
DB 478 GAACACAGGGGGCGGATACCGCTGCAAGTTGAGACCCCGCTCCGAAACAGCAATCATT 537
QY 154 ValProTyArgAerAerGlu1AlaH1sLeuArg11eMetLeuH1sAerLeuH1sSerLeu 173
DB 538 GTGCTCATGTCGCCGAGAGCACACCTGCGCTCTCTTACCACTGACCCCTTC 597

```

174 LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGlnValAlaAsnGlnThr 193
 598 TTGACGCCCGACAGCTTCTATGCGATCTATGATCCACAGCGCTGGAAATGAAACA 657
 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTrp 213
 658 TTTAAACAGGCGCAAACTGTGAACTGGGGGCGAGAGCGCTCGTGAATGAAAGTGG 717
 214 GlnCysPheIlePheIleAspValAspLeuProGlnAspAspArgAsnLeuTyrThr 233
 718 GACTGCTCTTCTTTCACGATGTGACCTCTTCCAGAAATGACACACATCTGTATGTG 777
 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu 252
 778 TGTGACCCCGGGGACCCCGCATGTGCGCTGCTATGAACATTTGGATACAGCCTC 837
 253 ProTyrSerAlaIlePheGlyLysValSerAlaLeuThrLysAspHisLeuLysIle 272
 838 CCGTACCCCGACAGTACTTGGAGAGAGTCTCAGACCTTACTCCTGACACAGTAAAGATG 897
 273 AsnGlyPheSerAsnAspPheTyrProGlyTyrGlyLysAspAspLeuAlaThrArg 292
 898 AATGCTTCCCAAGAAATACCTGGGCTGGGGTGGAGATGACGACTT-GCTACCAAG 956
 293 ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet 312
 957 GTGCGCCCTGGCTGGAGTGAAGATCTCTGGCCCCCACAATCTGTGAGACATATTAAGATG 1016
 313 IleLysHisSerThrGlnAlaThrAsnProValAsnLysCysArgTyrLysIleMetGly 332
 1017 GTGAAGACACCGAGATTAAGGCAATGAGAAATCCCAAGATTGACTCTCTGGTTC 1076
 333 GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn 352
 1077 GGTACCCAGAAATCTCTGACGAGATGGAGTGAACCTGACATACATACAGTTGCTGCT 1136
 353 LeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
 1137 CGAGAGCTGGGGCTCTTTATTCACAACTACAGACAGACATT 1178
 1137 CGAGAGCTGGGGCTCTTTATTCACAACTACAGACAGACATT 1178
 RESULT 7
 ABQ92036
 ID ABQ92036 standard; cDNA; 1853 BP.
 XX
 AC ABQ92036;
 XX
 XX 04-OCT-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 33.
 XX
 XX Human, cytostatic; antirheumatic; antiarthritic; vulnerary; analgesic;
 KM antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KM neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KM anticancer; fungicide; antidiabetic; antisthmatic; antiallergic;
 KM immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KM cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KM stem cell; growth factor; nervous system disease; neuropathy;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KM osteoporosis; severe combined immunodeficiency; SCID; infection;
 KM multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002065394-A1.
 XX
 PD 30-MAY-2002.
 XX
 PF 22-DEC-2000; 2000US-00745763.
 XX
 PR 18-MAR-1998; 98US-00040963.
 XX
 PA (JACOBS/) JACOBS K.
 PA (MCCOY/) MCCOY J M.

PA (LAWA/) LAWALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREAC/) TREACY M.
 PA (SPAUD/) SPAULDING V.
 PI Jacobs K, McCoy JM, Lawallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 DR WPI: 2002-582343/62.
 DR P-PSDB; ABP61819.
 XX
 XX Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune disease, bone disorders and lung or liver fibrosis.
 XX
 XX Claim 111, Page 144-145; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention
 XX
 XX SQ Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;
 Alignment Scores:
 Pred. No.: 6,79e-68 Length: 1853
 Score: 684.00 Matches: 133
 Percent Similarity: 63.87% Conservative: 42
 Best Local Similarity: 48.54% Mismatches: 97
 Query Match: 33.68% Indels: 3
 DB: 6 Gaps: 2
 US-10-661-430-1 (1-383) x ABQ92036 (1-1853)
 QY 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProGlnHisLeuValGlyProIle 114
 DB 358 CCAGCTCTCTCAAGGCTGCTCCCTACTGTCAGAAAGATCTCTCTTGTGGTCTGTG 417
 QY 115 ArgVal---PheLeuAspGluProAspPheLysThrLeuGlnLysIleTyrProAspThr 133
 DB 418 TCGGTGCTCTTACGCCAGTGCATCTGCGACAGATGTTGGACGGAAATCCCGGGTA 477
 QY 134 HisAlaGlyGlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153


```
DB 478 GAACCAAGGGGCGGATACCGCCTTGACGATTGTGAGCCCGCTCCGAAACGACCATCAT 537
QY 154 ValProTyrArgAspArgGluAlaHisLeuArg1LleMetLeuHisBAenLeuHisSerLeu 173
DB 538 GTGCTCTATCTGCTCCGGAGACCACTGCGCTCTTACCACTGACCCCTTC 557
QY 174 LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAenGlnThr 193
DB 598 TTGCACGCGCAGACGCTGCTTATGGCATCTATGTATCATCCACAGGCTGTAATGAAACA 657
QY 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyr 213
DB 658 TTATAACAGGGCAAAACGTGTGAACGTGGGGGTGCGAGAGCCCTCGGTGAAGAAGTGG 717
QY 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr 233
DB 718 GACTGCTGTTCTTGACGATGTGACCTTGTGCAAAATATGACCAATCTGATGTG 777
QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu 252
DB 778 TGTGACCCCGGGGAGACCCCGCATGTTGCCGTTGTATGAACAAGTTTGATACAGCTTC 837
QY 253 ProTyrSerAlaIlePheGlyLysIleSerAlaLeuThrLysAspHisLeuLysIle 272
DB 838 CCGTACCCCGCAGTACTCGAAGAGTCTCAGCACTTACTCTGACCACTGACCTGAAGATG 897
QY 273 AsnGlyPheSerAsnAspPheTyrGlyTyrGlyGlyValAspAspAspLeuAlaThrArg 292
DB 898 AATGGCTTCCCAATGATATCTGGGGCTGGGGGTGGTGGATGACGACTT-GCTTACCAAG 956
QY 293 ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet 312
DB 957 GTGCGCTGCTGGATGAAAGATCTCTGCGCCCGCATCTGTAGAGACACTTAAGATG 1016
QY 313 IleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGly 332
DB 1017 GTGAAGACCGAAGAGATTAAGGGCATGAAGAAATCCCAAGATTGTGACCTTCGTGTC 1076
QY 333 GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn 352
DB 1077 CGTACCAGAAATCTCTGACGCAAGATGGATGAACCTGACATATCCAGTTGCTGGCT 1136
QY 353 LeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
DB 1137 CGAGAGCTGGGGCTCTTTATACCAATCAATCAAGACACATTT 1178
RESULT 8
ABLI4208/c
ID ABLI4208 standard; cDNA; 4093 BP.
XX
XX ABLI4208;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37106.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX
XX MO2001.71042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter uC, Adams M, Li PWD, Myers EW;
```

```
XX
XX WPI; 2001-656860/75.
DR
XX P-PSDB; ABB70105.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 37106; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 4093 BP; 1129 A; 905 C; 927 G; 1132 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2,38e-66 Length: 4093
Score: 675.00 Matches: 161
Percent Similarity: 50.00% Conservative: 59
Best Local Similarity: 36.59% Mismatches: 137
Query Match: 33.23% Indels: 83
DB: Gaps: 11
US-10-661-430-1 (1-383) x ABLI4208 (1-4093)
QY 5 HisLeuAlaValAlaArgLeuLysSerLeuLeuValLeuCysAlaValLeuLeuVal 24
DB 2852 CATATTGCTTGGCAGAGGATATCTTATCAATTATTTTGGAAACGATATTAGCTT-- 2795
QY 24 HisAlaMetIleTyrLysIleProSerLeuTyrGlnAsnLeuThr-----IleG1 41
DB 2794 ----CGGCGGTGCACAAAGATGCTCATATCTACGGAACGTAAGCAGCAGATGAGCCGG 2739
QY 41 ySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrLase 61
DB 2738 AGCGACGTGA--GCATCCAGGCTGCCCTTCCCGCTCGCTTCAAAAGACAGAGA 2682
QY 61 rThrSerAspAspLeuAspThrTrpAsnSerThrPhe---SerProIleSerGluVal 80
DB 2681 GCGGACCAAGAGCTAATGCGGACCCCACTTACATTAAGAACTGTGATTCACCGGC 2622
QY 80 IaenGlnThrSerPheMetGluAspIleArgProIleLeu-----Ph 94
DB 2621 AATCTTACTTTCATTCACAAAGACTTAACGGCTTCCTGCTGGGCAAAAGATTTT 2562
QY 94 eProAspAsnGln-----ThrLeuGlnPheCysAsnGln----- 105
DB 2561 GCCCGCGCAGAAATCCACATCCGCTCTTGGCACTGCATGCATCCGATCCGATCCCG 2502
QY 106 -----ThrProHisLe 110
DB 2501 TGAATGTAAGCTGGCTGGCTTCTTTAGACAGATTTAAATACGTTCACTTGT 2442
QY 110 uValGlyProIleArgValPheLeuAspGluProAspPheLysThrLeuGlu---LysI1 129
DB 2441 AGGTGAGCCCATACCGCCAAACAGACAGACAGTGCATCGAGATTAATGAGCGGAGCT 2382
QY 129 eTyrProAspThrHisAlaGlyLysIleValMetProLysAspCysValAlaArgHisSar 149
DB 2381 TGGACCTCTTTTGGCGCTGTGGCTTTCGAGCTGAAAGCTGCAATGCCAGCATCA 2322
QY 149 gValAlaIleIleValProLysArgAspArgGluAlaHisLeuArgIleMetLeuHisAs 169
DB 2321 CGTGGCTATTGTGTGCTTCCGCGATGATAGCGCCATCTATTCTTTCTTCCGGGCAA 2262
```



```
QY 169 nleuHisSerLeuLeuAlaIyGInGInLeuAspTyrAlaIlePheIleValGluGlnVa 189
DB 2261 CATCCACCACCATTTCTGATGAGACAGCGCATCGCTATGSCATTTTCATTGTGAGAGAGAC 2202
QY 189 lAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerAr 209
DB 2201 CAACGGGAAGCCCTTAAATCGGGCTGCGCATGATGAACATGTGTTATTTGGAGCGCTTAA 2142
QY 209 GHeuTyrProTPrGInCysPheIlePheHisAspValAspLeuLeuProGluAspAspAr 229
DB 2141 GCTGTAACCACTGGGATGTTTATATATCCAGATGTCGATCTTGCCCTTGGACGACCG 2082
QY 229 GAsnLeuTyrThrCysProIleGInProArgHisMetSerValAlaIleAspLysPheAs 249
DB 2081 CAATCTTTCAACTGTGCCACGTGACCGCGGACACATGTCATGCGCTATGACACCGTGA 2022
QY 249 n-----Ty 250
DB 2021 CTTTCAGCGAGTAGAGCAAGAAGATMAACAATCAATGTAATGATTCAATTTCTTTA 1962
QY 250 rlyLeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLys 270
DB 1961 TAGGTTGCTTATCCATCAATATTTGGAGGTGTTCGCAATGACGCGGACGACCTTTCA 1902
QY 270 slySileAsnGlyPheSerAsnAspPheTPrGlyTyrGlyGlyLysAspAspLeuAl 290
DB 1901 GGGCCGTAATGGGCTTCTCAAACTCGTTCTTGCGCGCGCGGACGATGACGACATGTC 1842
QY 290 aThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTy 310
DB 1841 CAACGAGGTGAAGCAACGCAACCTATTCATATCAAGGTATCCGGTCAACATAGCTCGCTA 1782
QY 310 rlyMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrIleAl 330
DB 1781 CAAGATGCTGAAGCATCAGAGAGAAAGGCCAATCTTAAGCGGTCACTGACCATATCTAT 1722
QY 330 eMet----- 331
DB 1721 TATCCATATATTAGACTMAAANATGAAATTTTACCTTTCTTTTGGACAGCTATGA 1662
QY 332 -----GlyGlnThrLysArgArgTPrThrArgAspGlyLeuSerAsnLeuLys 347
DB 1661 AACTTACAGAAATGGCATGACGCAAAATAGAA-----CAGATGAGATCAACTGCAATA 1608
QY 347 sTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
DB 1607 GTATTCTATCTACAGCATCAAACTTTCAACTTTCACCTTGATCTTAGCAGAGCTA 1550
RESULT 9
AA066892
ID AA066892 standard; cDNA; 2304 BP.
AC AA066892;
XX
XX 25-MAR-2003 (revised)
DT 18-NOV-1994 (first entry)
XX
XX Glycosyltransferase hybrid.
XX
XX Galactosyltransferase; sialyltransferase; hybrid protein;
KM glycosyltransferase; glycoprotein; glycolipid; oligosaccharide; YEPGSTB;
KM glycosylation; Saccharomyces cerevisiae; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..2304
FT /tag= a
FT /product= "galactosyltransferase-sialyltransferase hybrid
FT protein."
XX
XX W09412646-A1.
```

```
PD 09-JUN-1994.
XX
XX PF 15-NOV-1993; 93WO-BP003194.
XX
XX PR 27-NOV-1992; 92EP-00810924.
XX
XX PA (CIBA ) CIBA GEIGY AG.
XX
XX PI Berger EG, Matzele M, Iwanow SX;
XX
XX DR WPI, 1994-200274/24.
XX
XX DR P-PSDB; AAR55709.
XX
XX PT Proteins with glycosyl transferase activity - useful for synthesis or
XX modification of glyco-proteins, glyco-lipid(s) and oligosaccharide(s).
XX
XX PS Disclosure; Page 52-57; 67pp; English.
XX
XX CC Hybrid glycosyltransferases (see also AAR55708) consisting of a membrane-
XX bound galactosyltransferase linked at its C-terminal to the N-terminal of
XX a soluble sialyltransferase have been expressed in Saccharomyces
XX cerevisiae using plasmid YEPGSTA. (Updated on 25-MAR-2003 to correct PN
XX field.)
SQ
SQ Sequence 2304 BP; 552 A; 658 C; 576 G; 518 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.69e-66 Length: 2304
Score: 668.50 Matches: 140
Percent Similarity: 54.18% Conservative: 74
Best Local Similarity: 35.44% Mismatches: 132
Query Match: 32.91% Indels: 49
DB: Gaps: 6
US-10-661-430-1 (1-383) x AA066892 (1-2304)
QY 14 LeuLeuValLeuCysValAlaValLeuLeuValHisAlaMetIleTyr----- 29
DB 76 CTCGTGGCGCGCTGCGCTCTGACCTTGCGGTCAACCTGTTACTACTGCTGCGCGC 135
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
DB 136 GACCTGAGCCCGCTGCGCCCACTG-----GTGCGAGTCTCCACACCGCTG 180
QY 47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
DB 181 CAGGCGGCTCGAACAAGTCCGCGCCGATCGGGAGTCTCGGGAGCTCCGACCGGA 240
QY 61 -----SerThrSer 63
DB 241 GGGGCGCGCGCGCGCTCTAGGCGCTCTCCAGCGCGCGCGCGGATGCGCATTC 300
QY 64 AspAspLeuLeuAspThrTPrAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
DB 301 AGCCCAAGTCGTGATTTCTGCGCGCTGCGCGCGCTGACCACTTGACTCGGTC----- 351
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
DB 352 -----CAspTgCCCAACACACGACCGGACCTGCTGCGCGCTGCG 390
QY 104 AsnGlnThrProProHisIleLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
DB 391 CCTGAGAGTCCCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 450
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyGlyHisGlyMetProLys 142
DB 451 CTGAGAGCTCGTGGCAAAAGCAGAACCAATGGAAGATGGCGCGCGCGCGCGCGCG 510
QY 143 AspCysValAlaAsnHisArgValAlaIleIleValAlaProTyrArgAspArgGluAlaHis 162
DB 511 GACTGCGTCTCTCTCAAGAGTGGCGCATTCATTCATTCATTCATTCATTCATTCATTCAT 570
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIyGInGInLeuAspTyrAla 182
```

```

Db      571 CTCAGACTAGCGCTATATTATTTGGACCCAGCTCTGAGCGGACGAGCTGACTATGGC 630
Qy      183 ILePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyValMetLeuVal 202
Db      631 ATCTATGTTATCAACACGCGGAGACACTATATTCATATGCTCTAATGCTCAATGTT 690
Qy      203 GLYTYrAspValAlaSerArgLeuTYrProTPrGlnCysPheIlePheHisAspValAsp 222
Db      691 GGTTCACAGACCTTGGAAGACTATGACTACACCTGCTTGTGTTAGTACAGCGGAC 750
Qy      223 LeuLeuProGluAspAspArgAsnLeuTYrThrCysProIleGlnProArgHisMetSer 242
Db      751 CTCATTCATGATGATGACCATTAATGCGTACAGGTGTTTTCACAGCCACGCGCATTTCC 810
Qy      243 ValAlaIleAspIysPheAsnTYrIysLeuProTYrSerAlaIlePheGlyIleSer 262
Db      811 GTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 870
Qy      263 AlaLeuThrIysAspHisLeuIysIleAsnGlyPheSerAspAspPheTPrGlyTPr 282
Db      871 GCTCTAGTAAACAACGTTTCTTAACCATGATGATGATGATGATGATGATGATGATG 930
Qy      283 GLYGLYGLuAspAspAspLeuValThrArgThrSerMetAlaGlyLeuIysValSerArg 302
Db      931 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 990
Qy      303 TYrProThrGlnIleAlaArgTYrIysMetIleIleHisSerThrGluAlaThrAspPro 322
Db      991 CCAATGCTGTGTGCGGAGGTGTGCGATGATGCGCATCAAGACCAAGAAAGAAAGAA 1050
Qy      323 ValAsnIysCysArgTYrIysIleMetGlyGlnThrIysArgArgTPrThrArgAspGly 342
Db      1051 CCCAATCCTCAGAGCTTGAACCGAATGACACACAAAGAGAGACAAATGCTCTGATGCT 1110
Qy      343 LeuSerAsnLeuIysTYrIysLeuValAsnLeuGlnLeuIysProLeuTYrThrArgAla 362
Db      1111 TTGAACCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1170
Qy      363 ValValAspLeuLeuGluIysAspCysArgArgGluLeuIysArgArg 377
Db      1171 ACAATGACATC-----GGGACACGAGCTAGATCCGTCGA 1206

RESULT 10
AA003464 standard; DNA; 2615 BP.
XX
AC      AA003464;
XX
DT      25-MAR-2003 (revised)
DT      15-AUG-1990 (first entry)
XX
DE      cDNA encoding N-acetylglucosamine (beta 1-4) galactosyl transferase.
XX
KW      Saccharide; teratocarcinoma F9 cell strain;
KM      N-acetylglucosamine (beta 1-4) galactosyl transferase; ss.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      CDS 73..1269
FT      /tag= a
XX
PN      JF02027987-A.
XX
XX      30-JAN-1990.
XX
PF      15-JUL-1988; 88JP-00176346.
XX
PR      15-JUL-1988; 88JP-00176346.
XX
PA      (GOKK-) GOKKO HOJIN.
XX

```

```

DR      MPI; 1990-072976/10.
DR      P-PSDB; AAR05674.
XX
PT      N-acetyl-glucosamine (beta 1-4) galactosyl transferase coding c-DNA -
XX      useful for providing the transferase and hence for saccharide synthesis.
XX
XX      Disclosure; Fig 2; 9pp; Japanese.
XX
PS      The cDNA is very useful for prodn. of N-acetylglucosamine (beta 1-4)
CC      galactosyl transferase for in vitro synthesis of saccharide chains. See
CC      also AA003463. (Updated on 25-MAR-2003 to correct PD field.)
XX
SQ      Sequence 2615 BP; 572 A; 694 C; 645 G; 704 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,84e-66 Length: 2615
Score: 668.00 Matches: 117
Percent Similarity: 54.69% Conservative: 73
Best Local Similarity: 35.68% Mismatches: 128
Query Match: 32.89% Indels: 46
DB: 2 Gaps: 5

US-10-661-430-1 (1-383) x AA003464 (1-2615)
Qy      14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTYr----- 29
Db      148 CTCGCGCGCGCTGCGCTGCACTTCCGCTCACCCCTGTTACTACCTGCGCGCGC 207
Qy      30 -----LysIleProSerLeuTYrGluAsnLeuThrIleGlySerThrLeuIle 46
Db      208 GACCTGAGCCGCGCTGCCCACTG-----GTGGAGCTTCACACCGCTG 252
Qy      47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
Db      253 CAGGGCGGCTCAACAAGTCCGCCGCGCATCGGCACTCCTCGGGAGAGCTCGGACCGGA 312
Qy      61 -----SerThrSer 63
Db      313 GGGGCGCGCGCGCTCCTCTAGCGCGCTCCTCCAGCGCGCGCGCGCGCGCGCTCC 372
Qy      64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAlaAsnGlnThr 83
Db      373 AGCCAGCTGCTGATCTGCGCCCTGCGCCGCTAGCACTGATGATGATGATGATGATG 423
Qy      84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuIysPheCys 103
Db      424 -----CCAGTCCCGCCACACACCGCACTGCTGCGCGCGCTGC 462
Qy      104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro---Asp 122
Db      463 CTTGAGGAGTCCCGCTGCTGCGCGCGCTGATGATGATGATGATGATGATGATGATG 522
Qy      123 PheLeuThrLeuGlnIysIleTYrProAspThrHisAlaGlyIysHisGlyMetProIys 142
Db      523 CTGAGCTCCTGCGCAAGACGAAACCAATGTGAAGATGGCGCGCGCTATGCGCCGAGG 582
Qy      143 AspCysValAlaIleArgHisArgValAlaIleIleValIleProTYrArgAspArgGluAlaHis 162
Db      583 GACTGCGCTCTCTCTCAACAAGTGGCCATCATCTCAATTCGCGCAACCGCGAGAGAC 642
Qy      163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuValIysGlnGlnLeuAspTYrAla 182
Db      643 CTCAGACTAGCTGATATATTATTGACACCCAGCTCGAGCGCGCGAGAGCTGAGCTATGCG 702
Qy      183 ILePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyIysLeuMetLeuVal 202
Db      703 ATCTATGTTATCAACACGCGGAGACACTATATTCATATGCTCTAATGCTCAATGTT 762
Qy      203 GLYTYrAspValAlaSerArgLeuTYrProTPrGlnCysPheIlePheHisAspValAsp 222
Db      763 GGTTCACAGACCTTGGAAGACTATGACTACACCTGCTTGTGTTAGTACAGCGGAC 822
Qy      223 LeuLeuProGluAspAspArgAsnLeuTYrThrCysProIleGlnProArgHisMetSer 242

```

```

DB      823 CTCATTCGAATGAATGATCATATGCGTACAGGTGTTTTTTCACAGCCACGACATTTCC 882
QY      243 VALAIAIIEAPLYPHEASNTYRYSLEUPROTYRSEALIIIEPHEGLYGLYIESER 262
DB      883 GTTGCAATGGAATGATTTGGATTTGATCCCTAATCTTATTTTCAATTTTGGAGGTGTCCT 942
QY      263 AIAIEUTHLYEASPHISLEULYSYLIEANGLYPHESEASPHETRPGLYTRP 282
DB      943 GCTCTAAGTAAACACACAGTTTCTAACCATTCGATTTCCATATATATTATGGGGCTG 1002
QY      283 GYGLYGLIUAAPSPAPSPLEUALATHRARGTTHSRMEEALIGLYLEULYSVALSERARG 302
DB      1003 GGAGGAGAAGATGATGACATTTTAAACAGATTAGTTTAAAGGACATGCTATATCTCGC 1062
QY      303 TYRPROTHGINLIEALARGTYRYSMETLIEYSHISERTHGLUALATHRANPRO 322
DB      1063 CCAAAATGCTGTGGTGGGAGGTGGCATGATCCGCACCTCAAGACAAAGAAAATGAA 1122
QY      323 VALAENLYCYEARGTYRYSILEMETGLYINTHLYEARGTYRTHRARGAPGLY 342
DB      1123 CCCAATCTCTCAGAGGTTTGACCGAATGACACACAAAGACAAATGCTCTCTGATGT 1182
QY      343 LEUSERASNLLEULYSYRYSLEULVALSENLEUGIULEULYSPROLEUTYRTHRARGALA 362
DB      1183 TTGAATCACTACCTACCTACAGGTGATGTACAGAGATACCCATTTGATATCCCAATTC 1242
QY      363 VALVALAPLEU 366
DB      1243 ACAGTGACATTC 1254

```

RESULT 11

AA003463
ID AA003463 standard, cDNA; 3923 BP.

AC AA003463;

XX 25-MAR-2003 (revised)
DT 15-AUG-1990 (first entry)

XX cDNA encoding N-acetylglucosamine (beta 1-4) galactosyl transferase.

XX Saccharide; teratocarcinoma F9 cell strain;
KM N-acetylglucosamine (beta 1-4) galactosyl transferase; ss.

XX Mus musculus.

OS Mus musculus.

FT CDS

FT Location/Qualifiers

FT /*tag= a

FT 2073..2078

FT /*tag= b

FT

XX JP02027987-A.

XX 30-JAN-1990.

XX 15-JUL-1988; 88JP-00176346.

XX 15-JUL-1988; 88JP-00176346.

XX (GOKK-) GOKKO HOJIN.

XX WPI, 1990-072976/10.

XX P-PSDB; AAR05673.

XX N-acetyl-glucosamine (beta 1-4) galactosyl transferase coding c-DNA -

XX useful for providing the transferase and hence for saccharide synthesis.

XX Claim 1; Page 1; 9pp; Japanese.

XX Sequence was determined from cDNA prep'd. from mRNA isolated from mouse

XX teratocarcinoma F9 cell strain. The cDNA is very useful for prodn. of N-

CC acetylglucosamine (beta 1-4) galactosyl transferase for in vitro synthesis
CC of saccharide chains. See also AA003464. (Updated on 25-MAR-2003 to
CC correct pp field.)

XX SQ Sequence 3923 BP; 831 A; 1010 C; 970 G; 1112 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,63e-65	Length:	3923
Score:	667.50	Matches:	123
Percent Similarity:	65.06%	Conservative:	52
Best Local Similarity:	45.72%	Mismatches:	93
Query Match:	32.87%	Indels:	1
DB:	2	Gaps:	1

US-10-661-430-1 (1-383) x AA003463 (1-3923)

```

QY      99 ThrleuGlnPheCYEANGlNThrProBhoiSleuValIGlyProIleArgVal--Phe 117
DB      525 TCGCTGCCAGCTTCCCTGAGAGATCCCGCTGCTCGTTGGCCCATGCTGATTCATCTT 584
QY      118 LeuAPGluProAPRPhelyThleuGluLysIleTyRProAPRThiSAlIGly 137
DB      585 AATATTGCTGTGGATCTGAGCTTTTGGCAAGAAACCAAGAGATAAAGCGGCGG 644
QY      138 HisGlyMeCProlYsAPCYEValAlaArgHisArgValAlaIleIleValProTyArg 157
DB      645 CGTACTCCCCCAAGAGATGTGTCTCTCTCTCAAGAGTGGCCATCATCATCCATTCGCT 704
QY      158 APARGluAlaHisIleuArgIleMetLeuHisIleuLeuHisIleuLeuAlaLysGln 177
DB      705 AACCGGACAGAGACATCTCAAAATACCTGCTGATTTATTTGATATCCATCTTCAGCGCAG 764
QY      178 GlnLeuAPRYRAlaIlePheIleValIGluGlnValAlaANGlNThrPhasnaArgly 197
DB      765 CAATCGACTAAGGATTTACGATCATCATAGCTGAGACACCATGATCATGAGACT 824
QY      198 LysLeuMeaenValIGlyTyRAspValAlaSerArgLeuTyRProTRPGLNcysPheIle 217
DB      825 AAGCTGCTCAATATTGCGTTTCAAGAGGCTTGAAGAGCTATGATTTACATCTGTTTGTG 884
QY      218 PheHisAPRValAPLeuLeuProGluAPSPARGAenLeuTyRThcysProIleGln 237
DB      885 TTCAATGATGATGACCTCATTCGATGAGACGACGTAATGCTTACAGGTGTTTTCGAG 944
QY      238 ProArgHisMetSerValAlaIleAspLysPheAsnTyRYSLeuProTySerAlaIle 257
DB      945 CCACGGCACATTTCTGTTGCAATGACAAAGTTCCGGGTTTACCTTCATATGTTCAAGTAT 1004
QY      258 PheGlyGlyIleSerAlaLeuThrLysAspHisIleuLysIleANGlyPheSerAsn 277
DB      1005 TTTGAGAGTGTCTGCTCTCAAGTAAACACAGTTTGTGCACTAAATGCAATCCCTAAT 1064
QY      278 APPhetRPGIlyTRPGIlyGluAPSPAPSPLEUALATHRARGTTHSRMEEALIGLY 297
DB      1065 AATTATTGGGCTGGGAGAGAGATGACGACATTTTAAACAGATTAGTTCATTAAGGC 1124
QY      298 LeuLysValSerArgTyRProThGINLIEALARGTYRYSMETLIEYSHISERTHR 317
DB      1125 ATGCTATATACGCTCAATGATCTGTAGTGAAGAGGTGTCAAATGATCCGATTTCAAGA 1184
QY      318 GluAlaThranProValAsnLysCYEARGTYRYSILEMETGLYINTHLYEARGARG 337
DB      1185 GACAAGAAAATGAGCCCAATCTCAGAGGTTTGACCGGATCGCACATACAAAGAAACG 1244
QY      338 TrpThrARGAPGLYLeuSerAsnLeuLysIleuValAsnLeuGluLeuLysPro 357
DB      1245 ATGGGCTTCGATGTTGTAACCTTACCTTACCAAGAGTGTGATGTACAGAGATACCGG 1304
QY      358 LeuTyRThrARGAlaValValAspLeu 366
DB      1305 TTATATACCAATTCACAGTGGACATTC 1331

```

RESULT 12

screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcino. Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) as a biochip; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcino. Associated (CA) gene copy number. In addition, the CA gene are useful as DNA vaccine and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA gene sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

CC Sequence 3953 BP; 832 A; 1018 C; 981 G; 1122 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1,65e-65	3953	123	52	93	1
Percent Similarity:	667.50					
Best Local Similarity:	65.06%					
Query Match:	45.72%					
	32.87%					

US-10-661-430-1 (1-383) x ACN45213 (1-3953)

99 ThrLeuGlnPheCyhAnglnThrProHlsLeuValGlyProIleArgVal---Phe 117
 551 TCGCTGCCAGCTTGGCTTACAGAGATCCCGCTGCTGTTGGCCCATGCTGATGACTTT 610
 118 LeuAspGluProAspPheLeuGlnValLeuTyProAspThrHisAlaGlyGly 137
 611 AATATTGCTGTGATCTGAGCTTTTGGCAAGAACCCAGAGATTAAGACGGCGGC 670
 138 HisGlyMetProLeuAspCyhValAlaArgHisArgValAlaIleValProTyArg 157
 671 CGTACTCCCGACAGACGTGTCTCTCCACAAAGTGGCATATCCCATTCCT 730
 158 AspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuValAlaGln 177
 731 AACCGGACGAGCATTTCAAAATACGTGATATTGATGATCCCATCTTACGGCCAG 790
 178 GlnLeuAspTyAlaIlePheIleValGlnValAlaAsnGlnThrPheAsnArgGly 197
 791 CAACCTGATATGAGCATCTACGTCATCATCAAGCTGAGACACCATGTTCAATGAGCT 850
 198 LysLeuMetAsnValGlyTyAspValAlaSerArgLeuTyProTrpGlnCysPheIle 217
 851 AAGCTGCTCAATATGCTTTCAGAGGCTTGAAGAGCATATGATTAACAACCTTGTG 910
 218 PheHisAspValAspLeuProGluAspAspAsnLeuTyThrCysProIleGln 237
 911 TTCAAGATGTGAGCTCATCTCCATGACGACGCGAAGCTCAAGGTGTTTTCGCGAG 970
 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyLysLeuProTySerAlaIle 257
 971 CCACGGCATATCTTGTGCAATGACCAAGTTCGGTTAAGCTGCAATATGTTACGAT 1030
 258 PheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAla 277
 1031 TTTCGAGGTCT 1090
 278 AspPheTrpGlyTyTrpGlyGluAspAspLeuAlaThrArgHisSerMetAlaGly 297
 1091 AATTATTGGGGTTGGGAGAGAGAGATGACGACATTTTAAACGATTAAGTTCAATAAGCC 1150
 298 LeuLysValSerArgTyProThrGlnIleAlaArgTyLysMetIleLysHisSerThr 317
 1151 AAGTCAATATCACTGCAATGCTGTAGTAGGAGGTGTCATATGATCCGATTCACAA 1210
 318 GluAlaThrAsnProValAsnLysCysArgTyLysIleMetGlyGlnThrLysArgArg 337
 1211 GACAAAGAAAATGAGCCCATCTCTCAGAGGTTTGAACGAGATGCAACATCAAGAAACG 1270

Qy 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyLysLeuValAsnLeuGluLeuLysPro 357
 Db 1271 ATGCCCTTCAGTGGTTTGAAGTCTTACCTACAGAGGTGAGATGATCAAGATACCCG 1330

Qy 358 LeuTyThrArgAlaValAlaAspLeu 366
 Db 1331 TTATATCCCAAAATCAGTCAGTGCATC 1357

RESULT 14

ID AAA53309 standard; cDNA, 1158 BP.

AC AAA53309;

DT 25-SEP-2000 (first entry)

DE Beta 1,4 galactose transferase (hGT) gene.

OS Beta 1,4 galactose transferase; hGT; transgenic plant; glycoprotein; ss.

OS Homo sapiens.

FN WO20034490-A1.

PD 15-JUN-2000.

PF 08-DEC-1999; 99WO-JP006881.

PR 09-DEC-1999; 98JP-00350584.

PA (SEKI/) SEKI T.

PA (FUJI/) FUJIYAMA K.

PI Seki T, Fujiyama K, Yoshida T;

DR WPI; 2000-431306/37.

DR P-PSDB; AAB03647.

PT Method for manufacturing a glycoprotein having human-type sugar chain, useful to produce glycoproteins for administration to animals and humans.

PS Example 1; Page 70-73; 79pp; English.

CC This sequence represents a human beta 1,4 Galactose transferase (hGT) gene. The gene is used in a method for manufacturing a glycoprotein with a human-type sugar chain. The method comprises obtaining a transformed plant cell, by introducing to a plant cell the glycosyltransferase gene and the gene of an exogenous glycoprotein, and cultivating the obtained transformed plant cell. Also included in the invention is a plant cell CC with a sugar chain adding mechanism which can conduct a transfer reaction of a galactose residue to a non-reducing terminal acetylglucosamine residue, where the sugar chain adding mechanism adds a sugar chain containing a core sugar chain and an outer sugar chain, where the core sugar chain comprises a number of mannose and acetylglucosamine, and CC where the outer sugar chain contains a terminal sugar chain portion with a non-reducing terminal galactose. The method can be used to produce a glycoprotein with a human-type sugar chain. The glycoproteins can be CC administered to animals, including humans. As the glycoprotein has a CC mammalian sugar chain (e.g. human type sugar chain) it is not antigenic as the glycosylation is of human type

CC Sequence 1158 BP; 251 A; 347 C; 296 G; 264 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	3.98e-66	1158	137	73	128	5
Percent Similarity:	666.00					
Best Local Similarity:	54.69%					
Query Match:	35.68%					
	32.79%					

US-10-661-430-1 (1-383) x AAA53309 (1-1158)

Qy	14	LeuLeuValLeuCyAlaValLeuLeuLeuValAlaMetIleTy-----	29
Db	37	CTCGGCGCGCTGCGCTTGCACCTTGCGACCTTGCGACCTTGCTTACCTGCGCGCGC	96
Qy	30	-----LysIleProSerLeuTyrgLuanLeuThrIleGlySerSerThrLeuIle	46
Db	97	GACCTGAGCGCGCTGCGCCCAACTG-----GTGGAGTCTCCACACCGCTG	141
Qy	47	AlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAla-----	60
Db	142	CAGGGCGGCTCGAACAAGTCGCGCGGCATCGGGAGAGTCTCGGGGAGTCCGAGCCGGA	201
Qy	61	-----SerHiser	63
Db	202	GGGGCCCGGCGCGCGCTCTGAGCGGCTCTCCAGCGCGCGCGGCTGGCGCATCC	261
Qy	64	AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr	83
Db	262	AGCCCACTGTGGATTCTGGCCCTGGCCCGCTAGCAATTGACTTCGGT-----	312
Qy	84	SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys	103
Db	313	-----CAAGGCCCAACACACGACGACGTGCGTGGCCGCGCTGC	351
Qy	104	AsnGlnThrProProHisiLeuValGlyProIleArgValPheLeuAspGluPro--Asp	122
Db	352	CCGAGAGATGCCCGCTACTAGGGGCCCACTGACTGAGTTAACTGCTCTGGAGC	411
Qy	123	PheLeuThrLeuGluLysIleTyrrProAspThrHisAlaGlyGlyHisGlyMetProLys	142
Db	412	CTGGAGCTGTGGCAAGACAAACCAATGTGAGATGGGGGCGCTATGCCCCCAAG	471
Qy	143	AspCysValAlaIaArgHisArgValAlaIleIleValProTyrrArgAspArgGluHis	162
Db	472	GACTGGTCTCTCTCCACAAGGGGCGCATCATTCATTCCATTCGGCAACCGGAGAGAC	531
Qy	163	LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrrAla	182
Db	532	CTCAAGTACTGGCTATATTATTATTGGACCCAGCTCTCCACGCGACGCTGACATATGCG	591
Qy	183	IlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnVal	202
Db	592	ATTATAGTTATCAACCAAGCGGAGACACTATTTCAATGCTCAAGCTCTCAATGTT	651
Qy	203	GlyTyrrAspValAlaSerArgLeuTyrrProTyrrGlnCysPheIlePheHisAspValAsp	222
Db	652	GCGTTTCAAGAACCTTTGAAGACATAAGACTAACCTGCTTGTGTTAGTACGCGGAC	711
Qy	223	LeuLeuProGluAspAspArgAsnLeuTyrrThrCysProIleGlnProArgHisMetSer	242
Db	712	CTCATTCATGATGATGACCATTAATGGGTACAGGTGTTTTTTCACAGCCACAGGCACATTGCC	771
Qy	243	ValAlaIleAspLysPheAsnTyrrLysLeuProTyrrSerAlaIlePheGlyGlyIleSer	262
Db	772	GTTCGATAGTATGATGATTGGATTGAGCTTACCTTAATGTTCCAGTATTTGGAGGTGCTCT	831
Qy	263	AlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAspAspPheTrpGlyTrp	282
Db	832	GCCTAAGTAAACAACAGTTTCTTAACCATCAAGATTCCTTAATAATTATGGGCTGG	891
Qy	283	GlyGlyGlyLysAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg	302
Db	892	GGAGGAGAAATGATGACATTTTAAACAGATTGATTTTGAAGGCATGTCCTATATCTCGC	951
Qy	303	TyrProThrGlnIleAlaArgTyrrLysMetIleLysHisSerThrGlnAlaThrAsnPro	322
Db	952	CCAAATGCTGTGTCGGAGAGTGTGCGATGATCGGCACCAAGACAAAGAAAATTGAA	1011
Qy	323	ValAsnLysCysArgTyrrLysIleMetGlyGlnThrTyrrArgArgTrpThrArgAspGly	342
Db	1012	CCCAATCTCAGAGGTTTGAACCGAATTGCACACAAAGAGCAATGCTCTGTAGTGGT	1071
Qy	343	LeuSerAsnLeuLysTyrrLysLeuValAsnLeuGluLeuLysProLeuTyrrThrArgAla	362

DB	Accession	Gene	Protein	Species	Length	Score	E-value	Annotations
QY	363	ValvA	AspLeu 366	ValvA	366	1072	1.1e-11	TTGACTCTACCTTACCAAGTCTGATGTACAGAGTACCATTTGATATCCCAATC 11
DB	1132	ACA	GTGACATC 1143	ACA	1143	1072	1.1e-11	TTGACTCTACCTTACCAAGTCTGATGTACAGAGTACCATTTGATATCCCAATC 11
RESULT 15								
ID	ABK90350	standard; DNA; 1158 BP.						
AC	ABK90350;							
DT	05-NOV-2002	(first entry)						
DE	cDNA encoding human beta 1-4 galactose transferase (hGT).							
GN	Gene; ss; beta1-4 galactose transferase; hGT; glycoprotein;							
KW	secretory production; sugar chain; acetylglucosamine; food; transgenic;							
XX	human.							
OS	Homo sapiens.							
XX	Synthetic.							
XX								
PH	Key	Location/Qualifiers						
FT	CDS	1..1158						
FT		/*tag= a						
FT		/product= "Beta 1-4 galactose transferase"						
FT		/EC number= "2.4.1.38"						
FT	mutation	replace(405,A)						
FT		/*tag= b						
FT	mutation	replace(408,A)						
FT		/*tag= c						
FT	conflict	replace(528,G)						
FT		/*tag= d						
FT	conflict	/citation= [OTHER]						
FT		replace(562,T)						
FT		/*tag= e						
FT	conflict	/citation= [OTHER]						
FT		replace(621..631,TN)						
FT		/*tag= f						
FT	conflict	/citation= [OTHER]						
FT		replace(1047,G)						
FT		/*tag= g						
FT		/citation= [OTHER]						
FT		/note= "OTHER= Biochem. Biophys. Res. Commun., 157, 657-663 (1988)."						
XX								
XX	WO200257468-A2.							
XX								
XX	25-JUL-2002.							
XX								
XX	18-JAN-2002; 2002WO-JP000361.							
XX								
XX	19-JAN-2001; 2001JP-00012519.							
XX								
XX	(FUJI/) FUJIYAMA K.							
XX	(SEKI/) SEKI T.							
XX								
XX	Fujiyama K, Seki T, Yoshida T;							
XX								
XX	WPI; 2002-590749/63.							
XX	P-PSDB; ABG31511.							
XX								
XX	Producing a glycoprotein with a human sugar chain comprises introducing							
XX	to a plant cell a gene of an enzyme capable of performing a transfer							
XX	reaction of a galactose residue, and a gene of a heterologous							
XX	glycoprotein.							
XX								
XX	Example; Page 77-80; 82pp; English.							
XX								
XX	This invention relates to a novel method for secretory production of a							
XX	glycoprotein with a human-type sugar chain. The method comprises							

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 21:00:21 ; Search time 1076 Seconds
(without alignments)
8853.460 Million cell updates/sec

Title: US-10-661-430-2
Perfect score: 1152
Sequence: 1 atgcttcgcattgtgc.....acttccaagtggttttag 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA_Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	100.0	1152	7	US-10-661-430-2
2	192.6	16.7	1839	10	US-11-097-143-18554
3	167.6	14.5	2002	3	US-09-925-301-230
4	159.6	13.9	1062	7	US-10-287-226-287
5	159.6	13.9	1201	7	US-10-287-226-285
6	159.6	13.9	3931	8	US-10-757-262-111
7	159.6	13.9	3931	8	US-10-768-158-5
8	157.2	13.6	1853	3	US-09-745-763-84
9	155.4	13.5	1864	3	US-09-989-722-235
10	155.4	13.5	1864	3	US-09-989-723-235
11	155.4	13.5	1864	3	US-09-989-729-235
12	155.4	13.5	1864	3	US-09-989-727-235
13	155.4	13.5	1864	3	US-09-989-731-235
14	155.4	13.5	1864	3	US-09-989-732-235
15	155.4	13.5	1864	3	US-09-991-073-235
16	155.4	13.5	1864	3	US-09-990-442-235
17	155.4	13.5	1864	3	US-09-991-163-235
18	155.4	13.5	1864	3	US-09-993-604-235
19	155.4	13.5	1864	3	US-09-990-456-235
20	155.4	13.5	1864	3	US-09-989-721-235
21	155.4	13.5	1864	3	US-09-992-598-235
22	155.4	13.5	1864	3	US-09-989-293A-235
23	155.4	13.5	1864	3	US-09-989-735-235

24	155.4	13.5	1964	3	US-09-990-444-235	Sequence 235, App
25	155.4	13.5	1964	3	US-09-991-181-235	Sequence 235, App
26	155.4	13.5	1964	3	US-09-989-730-235	Sequence 235, App
27	155.4	13.5	1964	3	US-09-990-436-235	Sequence 235, App
28	155.4	13.5	1964	3	US-09-993-687-235	Sequence 235, App
29	155.4	13.5	1964	3	US-09-989-734-235	Sequence 235, App
30	155.4	13.5	1964	3	US-09-997-653-235	Sequence 235, App
31	155.4	13.5	1964	3	US-09-989-724-235	Sequence 235, App
32	155.4	13.5	1964	3	US-09-989-728-235	Sequence 235, App
33	155.4	13.5	1964	3	US-09-990-441-235	Sequence 235, App
34	155.4	13.5	1964	3	US-09-993-667-235	Sequence 235, App
35	155.4	13.5	1964	3	US-09-997-428-235	Sequence 235, App
36	155.4	13.5	1964	3	US-09-997-666-235	Sequence 235, App
37	155.4	13.5	1964	3	US-09-990-438-235	Sequence 235, App
38	155.4	13.5	1964	3	US-09-990-562-235	Sequence 235, App
39	155.4	13.5	1964	3	US-09-990-711-235	Sequence 235, App
40	155.4	13.5	1964	3	US-09-989-726-235	Sequence 235, App
41	155.4	13.5	1964	3	US-09-998-156-235	Sequence 235, App
42	155.4	13.5	1964	3	US-09-990-437-235	Sequence 235, App
43	155.4	13.5	1964	3	US-09-991-157-235	Sequence 235, App
44	155.4	13.5	1964	3	US-09-997-514-235	Sequence 235, App
45	155.4	13.5	1964	3	US-09-997-573-235	Sequence 235, App

ALIGNMENTS

```
RESULT 1
US-10-661-430-2
; Sequence 2, Application US/10661430
; Publication No. US2004008695A1
; GENERAL INFORMATION:
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: Beta 1, 4-N-ACETYLGLACTOSAMINYLTRANSFERASES,
; FILE REFERENCE: 7148.001
; CURRENT APPLICATION NUMBER: US/10/661,430
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 60/411,242
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-661-430-2
```

```
Query Match 100.0%; Score 1152; DB 7; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTTTCGTCATTGGCAGTCGCCAGACTCAAGTCGTTGCTGTACTTTGCGGTT 60
Db 1 ATGGCTTTTCGTCATTGGCAGTCGCCAGACTCAAGTCGTTGCTGTACTTTGCGGTT 60
QY 61 CTTTCAATTAGTCATGCAATGATTTTATAAGATTCATGCTTTAAGAACTTACTATC 120
Db 61 CTTTCAATTAGTCATGCAATGATTTTATAAGATTCATGCTTTAAGAACTTACTATC 120
QY 121 GGCTCTCGACCCCTTATTCGACGTCGACGATGAGAGGAGCTGTCGGAATACGGT 180
Db 121 GGCTCTCGACCCCTTATTCGACGTCGACGATGAGAGGAGCTGTCGGAATACGGT 180
QY 181 TCCACTTCGATGATCTAATTGATACGTGAATTCACGTTTTCACCGATTTCTGAAGTT 240
Db 181 TCCACTTCGATGATCTAATTGATACGTGAATTCACGTTTTCACCGATTTCTGAAGTT 240
QY 241 AATCAGACTAGTTTATGAGAGACATTCGTCATTCCTGTCGGAACAGACTCTT 300
Db 241 AATCAGACTAGTTTATGAGAGACATTCGTCATTCCTGTCGGAACAGACTCTT 300
```

```

QY 301 CAATTGTATATGACACCTCCCACTGTCGACCCATCCGTTGATTTCTCGATGAG 360
DB 301 CATTTCTATATGACACCTCCCACTGTCGACCCATCCGTTGATTTCTCGATGAG 360
QY 361 CCCGACTTCAAACTCTCGAAGAAATCTATCCGGAACGCAAGCCGCTGACATGGAATG 420
DB 361 CCCGACTTCAAACTCTCGAAGAAATCTATCCGGAACGCAAGCCGCTGACATGGAATG 420
QY 421 CCTAAGATTTGTTGCAAGGACATCGGTTGCTATTATTTGCTGCTATAGATCGGAA 480
DB 421 CCTAAGATTTGTTGCAAGGACATCGGTTGCTATTATTTGCTGCTATAGATCGGAA 480
QY 481 GCACATTTGAGATAATGCTCCCAATTTGCACTGTTGCTGCGCAACCAATTTGAC 540
DB 481 GCACATTTGAGATAATGCTCCCAATTTGCACTGTTGCTGCGCAACCAATTTGAC 540
QY 541 TATGCAATTTTCAATTTGAGCAATGCGCAATTCAGCTTTTATGCGGGGAACTTAAG 600
DB 541 TATGCAATTTTCAATTTGAGCAATGCGCAATTCAGCTTTTATGCGGGGAACTTAAG 600
QY 601 AACGTTGATAGACGATGATGACGCTGACCAATGGAAGCTTCACTTTCATGAT 660
DB 601 AACGTTGATAGACGATGATGACGCTGACCAATGGAAGCTTCACTTTCATGAT 660
QY 661 GTGATTTTACGCGCGAAGATGACCGTACCTGTCACGTCGCAATTCACCAACGTCAT 720
DB 661 GTGATTTTACGCGCGAAGATGACCGTACCTGTCACGTCGCAATTCACCAACGTCAT 720
QY 721 ATGAGTGTAGCGATTCGATAATTCATTAATTTTCAATTTGCGGATCTTCCGCGGA 780
DB 721 ATGAGTGTAGCGATTCGATAATTCATTAATTTTCAATTTGCGGATCTTCCGCGGA 780
QY 781 ATGAGTGTAGCGATTCGATAATTCATTAATTTTCAATTTTCAATTTTCAATTTTGG 840
DB 781 ATGAGTGTAGCGATTCGATAATTCATTAATTTTCAATTTTCAATTTTCAATTTTGG 840
QY 841 GGTGGGCGCGAAGGACGACGATTTGGCGACGAGAACATGCTGCTGACCTGAAAGT 900
DB 841 GGTGGGCGCGAAGGACGACGATTTGGCGACGAGAACATGCTGCTGACCTGAAAGT 900
QY 901 TCAAGTATTCGACACCAATTTGCAATTAATTAATTTGAGCACTGCAAGGACGACG 960
DB 901 TCAAGTATTCGACACCAATTTGCAATTAATTAATTTGAGCACTGCAAGGACGACG 960
QY 961 AATCCAGTTATTAATGCGCTACCAAAATATGCGGCAAGGACGCGATGACGACG 1020
DB 961 AATCCAGTTATTAATGCGCTACCAAAATATGCGGCAAGGACGCGATGACGACG 1020
QY 1021 GACGCGCTAAGCAATCTGAATAGCTGTAATCTGTAATTTGAAGCTTCTTACAT 1080
DB 1021 GACGCGCTAAGCAATCTGAATAGCTGTAATCTGTAATTTGAAGCTTCTTACAT 1080
QY 1081 CGAGCGCTGCTGATTTGCTGCAAAAGATGCGCGGAGAGCTGCGAAGGACCTTTCA 1140
DB 1081 CGAGCGCTGCTGATTTGCTGCAAAAGATGCGCGGAGAGCTGCGAAGGACCTTTCA 1140
QY 1141 ACGGTTTATG 1152
DB 1141 ACGGTTTATG 1152

```

```

RESULT 2
US-11-097-143-18554
; Sequence 18554, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143

```

```

; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 4308
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18554
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-18554

Query Match 16.7%; Score 192.6; DB 10; Length 1839;
Best Local Similarity 58.7%; Pred. No. 3.6e-51;
Matches 333; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 399 GCACGCGGTGACATGATGATGCTTAAGATTTGTTGCAAGGACCTGTTCTTAT 458
DB 600 GCGCCCTGTTGCGGCTTGAAGCAATGCAATGCGCAATGCGGATGCTTATGT 659
QY 459 TGTGCTTATGAGATGCTGAAGCAATTTGAATATGCTCCACAAATTTGACCTGTT 518
DB 660 TGTGCTTATGAGATGCTGAAGCAATTTGAATATGCTCCACAAATTTGACCTGTT 518
QY 519 GCTGCGCAACCAATTTGATGATGATGCTTAATTTGATTTGAGCAAGTGGCAATGAC 578
DB 720 TGTGATGAGGAGCGGATGCGCTATGCAATTTGATTTGATTTGAGCAAGTGGCAATGAC 578
QY 579 GTTTAATGCGGGAATCTATGACCTTGAATGACGATGACGATGACGCTTACCAATG 638
DB 780 GTTTAATGCGGGAATCTATGACCTTGAATGACGATGACGATGACGCTTACCAATG 638
QY 639 GCAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
DB 840 GCAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
QY 699 GTGCTAATTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
DB 900 GTGCTAATTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
QY 759 ATTTTGGGATTTTGGGGAATCTATGATGATGATGATGATGATGATGATGATGAT 818
DB 960 ATTTTGGGATTTTGGGGAATCTATGATGATGATGATGATGATGATGATGATGAT 818
QY 819 TGAATTTGAAATGATTTTGGGTTGGGCGGAGAGGACGATTTGGCGAGAAC 878
DB 1020 TGAATTTGAAATGATTTTGGGTTGGGCGGAGAGGACGATTTGGCGAGAAC 878
QY 879 ATGATGCTGACTGAAGTTTCAAGATATCCGACCAATTTGCAAGATATTAATGAT 938
DB 1080 ATGATGCTGACTGAAGTTTCAAGATATCCGACCAATTTGCAAGATATTAATGAT 938
QY 939 TAAGCACTGACGGAAGCGAGATCC 965
DB 1140 TAAGCACTGACGGAAGCGAGATCC 1166

```

```

RESULT 3
US-09-925-301-230
; Sequence 230, Application US/09925301

```

Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 230
LENGTH: 2002
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-301-230

Query Match 14.5%; Score 167.6; DB 3; Length 2002;
Best Local Similarity 54.4%; Pred. No. 5,5e-43;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CCTAAGATTGTTGCAAGGCGATCGTGTCTATTATTATGCCCCATTAGAGATCGTGA 480
611 CCGAGGTTGTAGAGCCCGCTCCGAAACAGCCATCATTTGCTCATGTCGCCGGAG 670
481 GCACATTGAGAAATATGCTCCACAAATTGCACTGCTGCTGCCAACAACAAATTGAC 540
671 CACCACTGCGCGCTGCTCTACACCTGCACTCCCTCTTCCAGCGCCAGACGCTTGT 730
541 TATGCAATTTTTCATTTGAGGCAAGTGGCAATCAAGCTTATATGCGGGAACCTAATG 600
731 TATGCACTATATGATCATCAGAGCTGGAATGAACTTTAAACAGGCAAACTGTG 790
601 AAGTTGATACAGCTAGCATCAGCGCTTACCCATGCGAGCTTCACTTTTCAATAT 660
791 AACGTTGGGGTGCAGAGGCGCTGCGTATGAAGAGTGGAGCTGCTGTTTTCAGCAT 850
661 GTCCATTACTGCCCCGAAGATACCGTACCTGTACACTG---CAATTCAACACGT 717
851 GTGAGCTCTTGTGCGAAATATGACCAATCTGTATGTGTGACCCCGGGGACCCCGC 910
718 CATATGAGTGTAGCGATGATTAATTCAATTATACTTCAATTTGGCGATCTTGGC 777
911 CATGTGCGGTTGCTATGAACAAGTTTGGATACAGCTTCCGTAACCCAGATCTTGG 970
778 GGAATCAGTGCATACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTT 837
971 GGAATCTGAGCACTTACTCTGACGATCACTGAAGATGAATGGCTTCCCAATGATAC 1030
838 TGGGGTTGGGCGAGAGAGACGATTTGGGCGAGAAACATCATGCTGCTGACGAA 897
1031 TGGGGCTGGGGTGTGAGATACGACATTTGTACAGGGTGGCGCTGGCTGGGATGA 1090
898 GTTTCAGATATCCGACCAATTTGACGATTAATAATGATTAAGCACTGACGGAAGG 957
1091 ATCTCTGCGCCCCCAGATCTGTAGACACTTATAGATGTGAAGACCGAGAGATTAG 1150
958 ACGAATCCAGTTAATTAATGCGCTACAAATATATGCGCAAAAGCGCGATGACA 1017
1151 GGCATAGAGAAATCCCAAGATTGACCTCTGCTGCTGATACCAAGAAATCTGAGAG 1210
1018 CGTGACGCGCTTAAGCAATGTGAAGTCTGTAATCTGGAATTTGAAGCTCTTAC 1077
1211 CAAGATGGATGAATCACTGACATACCAAGTGTGGCTGAGAGCTGGGGCTCTTTAT 1270
1078 AC 1079
1271 AC 1272

RESULT 4

US-10-287-226-287
Sequence 287, Application US/10287226
Publication No. US2004008675A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsbrook, John P.,
APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ferenc,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eissen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangolli, Baha A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malyankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Patlurajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suren G.,
APPLICANT: Spadeina, Steven K.,
APPLICANT: Spylek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vernet, Corine A.M.,
APPLICANT: Zethusen, Bryan D.,
APPLICANT: Zhong, Wei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 287
LENGTH: 1062
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: CDS
LOCATION: (2)..(1030)
US-10-287-226-287

Query Match 13.9%; Score 159.6; DB 7; Length 1062;
Best Local Similarity 54.3%; Pred. No. 1.7e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

```

QY 389 ATCCGACACGCGCGGTGACATGGAATGCTTAAGATTTGTTGCAAGCATCTG 448
DB 297 ATTTAATATTGAGCCAGGGGGTCAATTGAGGCCAAGACTTAACCAACGATGGAAG 356
QY 449 TTGCTATTATTGTCCTTAATGAGATCTGTAAGCAATTTGAGAAATATGCTTCAAT 508
DB 357 TGGCAGTTCTCATCTTCTTCCGTAATCGCATGAACATCTTCAATTTTCTTACATC 416
QY 509 TGACCTGTTGCTGCGCAACACAAATTTGACATTTTCAATTTGAGACCAAGTGG 568
DB 417 TGATTCGAATGCTCCAGAACACGGGTGGAATTTGGTTTATGTCATTGAACAGACTG 476
QY 569 CGAATGACAGCTTTAATCGCGGAACTAATGAACGTGATCGAGTACATCAAGCC 628
DB 477 GCAACACACTTTTAAACCGTCATGCTTTCAATGTGGCTTCAAGAGCCATGAAG 536
QY 629 TCTACCATGACAGTGTCTTCAATCTTCAATGATGATTTTCTGCGGAAGTACCGTA 688
DB 537 ACAGTGTCTGAGCTGTGTAATCTTCCACGATGTGATCATCTAAGTAATGACGGA 596
QY 689 ACCTGTACAGTGTCTTCAATCAACAGTCAATGATGATGAGATGATTAATTTCAAT 748
DB 597 ACTATTAACGATGTGAGAAATGCAAGTCAATTTTGTGCAAGCTGATTAATACATGT 656
QY 749 ATAACTTCATATTGCGCGCATCTTGGCGGAATCAAGTGAATCAAAAGTCACTGA 808
DB 657 ATATTCTTCATTAAGAAATTTTGTGTGTGTAAGTGGCTGACAGTGAACAATTTA 716
QY 809 AGAAATCAATGATTTTGAATGATTTTGGGGTGGGCGGAGAGACGACGATTTGG 868
DB 717 GAAAGATCATGTTTCTTAATGCTTCTGGGGATGGGAGAGAAATGATGACCTTT 776
QY 869 CGACGACATGATGATGCTGATGAATTTTCAATATCCGACCAATTTGACGAT 928
DB 777 GGAACAGAGTCACTATGCTGATATATATGTAACCAACAGAGGAGCTTTGGAAT 836
QY 929 ATAAATGATTAAGCACTGACGAGACGACGAATTCAGTTAATTAATGCCGCTACAAA 988
DB 837 ACAAGTCAATTTCTCA-TCAACATAGAGTGAAGTCCAGTT--TTAGGACGTAATTAAT 893
QY 989 TAAATGGCCAAACGAGCGCGCATGACACGATGACGCTTAAGCAATCTGAAGTAAAGC 1048
DB 894 TACTAAGGATATCCAAAGAGCGTCAATGATGATGACGACAAATTTAATATAGGC 953
QY 1049 TCGTAAATCTGG 1060
DB 954 CAAAAATATCTGG 965

```

RESULT 5
US-10-287-226-285

Sequence 285, Application US/10287226
Publication No. US20040086875A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ference,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eilsen, Andrew,
APPLICANT: Ellerman, Karen,

```

APPLICANT: Gangoli, Esha A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malyankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark B.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderma, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vermet, Corine A.M.,
APPLICANT: Zethusen, Bryan D.,
APPLICANT: Zhong, Mei,
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remainder Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 285
LENGTH: 1201
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (24)..(1169)
US-10-287-226-285

```

Query Match 13.9%; Score 159.6; DB 7; Length 1201;
Best Local Similarity 54.3%; Pred. No. 1.7e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

```

QY 389 ATCCGACACGCGCGGTGACATGGAATGCTTAAGATTTGTTGCAAGCATCTG 448
DB 436 ATTTAATATTGAGCCAGGGGGTCAATTGAGGCCAAGACTTAACCAACGATGGAAG 495
QY 449 TTGCTATTATTGTCCTTAATGAGATCTGTAAGCAATTTGAGAAATATGCTTCAAT 508
DB 496 TGGCAGTTCTCATCTTCTTCCGTAATCGCATGAACATCTTCAATTTTCTTACATC 555

```

QY	509	TGCACTCGTTCCTCCGCAAAACAACAAATGAGCATGATCAATTTTCATTGTGGACAAGTGG	568
Db	556	TGATTCCAATGCTCCAGAGACGCGCTGGAAATTTGCTTTATGTCAATTGAACAGACTG	615
QY	569	CGAATCGACGTTTATTCGCGGAAACTVATGACGTTGATVATGACGTACACGCC	628
Db	616	GCACACAACCTTTTAAACCGTGGCATGCTTTTCAATGTGGGCTTCAAGAGGCCATGAAG	675
QY	629	TCTACCAATGACAGTCTTCATCTTTCAAGATGTGATTTTACGCCCGAAGATGACCGTA	688
Db	676	ACAGTGTCTGGACGTGTGAATCTTCAAGATGTGATCATCTCACTGAAAAGTACCGGA	735
QY	689	ACCTGTACACGTCGTCCAAATTCACCAACGTATATGATGTAGCGATCGATTAATTCATTT	748
Db	736	ACTATATACGATAGTGTGAGAAAATGCAACGTCAATTTGCTCCAAAGCTGATTAATACATGT	795
QY	749	ATPAACTTCATATTCGCGCATCTTTCGCGGAATCATGTCACACTAACAAAAGATCACCTGA	808
Db	796	ATATCTTCATATATTAAGAAATTTTGTGGTGTGAATGTGGGCTGACAGTGGAAACAATTTA	855
QY	809	AGAAAAATCAATGAGATTTTCGAATGATTTTGGGGCTTGGGGCGAGAGACAGCATTTGG	868
Db	856	GAAAGATTAATGTGTCTTCTTAATGCTTCTGGGGAATGGGAGAGAAATGATGACCTTT	915
QY	869	CGACGAGAACATCGATGCTGACCTGAAGATTTCAGATATCCGACACAATTCACAGAT	928
Db	916	GGAACAGAGTTTCAATATGTGTGATATTAATGAACAGACAGAGGAGACTTAGAAAAT	975
QY	929	ATPAAATGATTAAGCATCTCGACGAGACGACGATTCAGTTATTAATTCGCTACAAA	988
Db	976	ACAAGTCATTTCTCTCA-TCACCATAGAGGTGAAGTCCAGTT--TTTAGGACGGTATTAAT	1032
QY	989	TTAATGGCGCAACGAGAGCGCATGAGACAGTGAAGCGCCTAAGCATTCGAATATTAAGC	1048
Db	1033	TACTAAGTATTTTCAAGAGAGCTCAGTACATGATGACTGAAACAATTTATATATTAAGC	1092
QY	1049	TCGTAATCTCG	1060
Db	1093	CAAAATACTGG	1104
RESULT 6			
US-10-757-262-111			
Sequence 111, Application US/10757262			
Publication No. US20040197825A1			
GENERAL INFORMATION:			
APPLICANT: Karichecki, Venkateswarlu			
APPLICANT: Eliasof, Scott D.			
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING			
TITLE OF INVENTION: UROLOGICAL DISORDERS USING 4A390, 5A181, 211, 5687, 884,			
TITLE OF INVENTION: 1405, 636, 4421, 5810, 30905, 2045, 16405, 18560, 2047,			
TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,			
TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,			
TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 16816, 17827, 32620, 577, 619, 1423,			
TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,			
TITLE OF INVENTION: 2158, 8263, 15402, 16309, 16386, 21165, 30911, 41897, 1643,			
TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR			
FILE REFERENCE: MP103-007P1RNMONTM			
CURRENT APPLICATION NUMBER: US/10/757,262			
CURRENT FILING DATE: 2004-01-14			
PRIOR APPLICATION NUMBER: US 60/440,318			
PRIOR FILING DATE: 2003-01-15			
PRIOR APPLICATION NUMBER: US 60/444,783			
PRIOR FILING DATE: 2003-02-04			
PRIOR APPLICATION NUMBER: US 60/457,901			
PRIOR FILING DATE: 2003-03-27			
PRIOR APPLICATION NUMBER: US 60/468,775			
PRIOR FILING DATE: 2003-05-08			
PRIOR APPLICATION NUMBER: US 60/471,614			
PRIOR FILING DATE: 2003-05-19			

```

? PRIOR APPLICATION NUMBER: US 60/478,742
? PRIOR FILING DATE: 2003-06-16
? PRIOR APPLICATION NUMBER: US 60/488,529
? PRIOR FILING DATE: 2003-07-18
? PRIOR APPLICATION NUMBER: US 60/491,156
? PRIOR FILING DATE: 2003-07-30
? PRIOR APPLICATION NUMBER: US 60/499,594
? PRIOR FILING DATE: 2003-09-02
? PRIOR APPLICATION NUMBER: US 60/506,332
? PRIOR FILING DATE: 2003-09-26
? NUMBER OF SEQ ID NOS: 136
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 111
? LENGTH: 3931
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (298)...(1446)
? US-10-757-262-111

```

Query Match	13.9%;	Score 159.6;	DB 8;	Length 3931;
Best Local Similarity	54.3%;	Pred. No. 3.4e-40;		
Matches 365; Conservative	0;	Mismatches 304;	Indels 3;	Gaps 2;

QY	389	ATCCGACACGACCGCCGGTGGACATGGAATGGCTTAAGAAATGGTTGGCAAGGCATCGTG	448
Db	710	ATTTAGATATTGAGCCAGGGGGTTCATTGGAGGCCAANAAGACTGTAAACCAGATGGAAG	769
QY	449	TTGCTATTATTGGCCCTATAGAGATGTGTAGACACTTTGAGAAATATGCTCCACATT	508
Db	770	TGGCAGTTCCTATTCCTTCCGTAATGCGCATGAAACCTTCCAAATTTTTTCTTACATC	829
QY	509	TGCACCTGGTGTCCGCAAAACAACAAATTGGAACATATGCAATTTTTCATTGTGGAGCAATGG	568
Db	830	TGATTCCAATGCTCCAGAAACAGCGGTGGAAATTTGGCTTTATGTCAATTGAAACAGACTG	889
QY	569	CGAATCAGACGTTTAAATCGCGGGAACCTAAATGAACGTTGATACGACGTAGCATCACGCC	628
Db	890	GCACACAACTTTTAAACCGTGCATGCTTTTCAATGTGGGCTTCAAAGAGCCATGAAAG	949
QY	629	TCTAACCATGGCAGTGCTCATCTTTCATGATGTCATATGCGATTAAGTCGCCGAAGATGACCGTA	688
Db	950	ACAGTGTCTGGAGCTGTGTATCTTCCACAGATGTGATCATCTACCTGAATAATGACCGGA	1009
QY	689	ACCTGTACACGTGTCCAAATTCAAACAACGTATATGATGTAGCGATCGATTAATTCATT	748
Db	1010	ACTATTACGGATGTGGAAATGCAACGACTCATTTTGTGCAAAAGTCGGATTAATACATGT	1069
QY	749	ATTAACCTTCATATTCGGCGCATCTTGGCGGGAATCAGTGCATTAACAAAAGATCACTGTA	808
Db	1070	ATATTCTTCATATTAAGAAATTTTTTGGTGTGTATGTGGCTGCACAGTGAACAATTTTA	1129
QY	809	AGAAATCAATGATATTTTCAATGATATTTTGGGCTTGGGCGGAGAGACGACCATTTGCG	868
Db	1130	GAAAGATCAATGTTTTCTTAATGCTTCTGGGATGGGAGAGGAAGATGATGACCTTT	1189
QY	869	CGACGAGACATGAGTGGCTGGACTGAAGTTTCAGATATCCGACCAAAATTGCACGAT	928
Db	1190	GGAAACAGAGTTCACATATGCTGGAGATATATATATTAACCAACAGAGGAGACCTTAGAAAT	1249
QY	929	ATTAATAATGATTAAGACATCGACGGAACGACGCAATCCAGTTAATTAATATGCCGTACAAAA	988
Db	1250	ACAAGTCAATTTCTCTCA-TCAACATAGAGGTGAAGTCCAGTT--TTTAGACGGTATTAAT	1306
QY	989	TAAATGGCCAAACGAAGCGCCGATGCAACAGTACGGCTTAAGCAATCTGAAGTATTAAC	1048
Db	1307	TACTAAGGTATTCCAAGAGACGTCACATGATCATGATGAGCTGAACAATTTATATATATAGGC	1366
QY	1049	TGCTAAATCTGG	1060
Db	1367	CAAAATATCTGG	1378

```
RESULT 7
US-10-768-158-5
; Sequence 5, Application US/10768158
; Publication No. US20040204359A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Siles-Santiago, Immaculada
; APPLICANT: Karicheh, Venkateswulu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012P1R0NM01M
; CURRENT APPLICATION NUMBER: US/10/768,158
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3931
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (298) ... (1446)
US-10-768-158-5

Query Match 13.9%; Score 159.6; DB 8; Length 3931;
Best Local Similarity 54.3%; Pred. No. 3.4e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACAGCAGCGCGTGAATGAGTGAATGCTTGAAGGATGCTG 448
DB 710 ATTGATATTGAGCCAGGGGTCATTTGAGGCCAAGAGCTGAACCCAGATGAAG 769
QY 449 TTGCTATTATTTGCTTATTAAGATGCTGAAGCATTGAGATTAATGCTCAATT 508
DB 770 TGGCAGTTCTCATCTTCCGATGCGCATGAACATCTTCCAAATTTTCTTACATC 829
QY 509 TGCAGCTGTTGCTGCCAACAACAATTGACATGATTTTTCATTTGAGAGCAATG 568
DB 830 TGATTCGAATGCTCAGAGAGAGCGGCTGAATTTGCTTATGATGAAGAGACTG 889
QY 569 CGAATCAGAGCTTATTCGCGGGAATTAATGAAGCTTGAAGAGAGTCAAGCC 628
DB 890 GCACACACCTTTAACCCTGCGATGCTTTTCAATGCGGCTTCAAGAGGCCATGAAG 949
QY 629 TCAACCATGCGAGTCTTCACTTTTCATGATGTCGATTTACTGCCGAAGATGACCT 688
DB 950 ACAAGTCTGCGAGCTGTAATCTTCCACAGATGATCACTTCACTGAAGATGACCGGA 1009
QY 689 ACCGTACAGCTGTCATTAACACAGCTATATGATGATGATGATGATGATGATGAT 748
DB 1010 ACTATTAAGATGAGAGAAATGCGACGCTCAATTTTGTGCGAAGCTGATTAATAATG 1069
QY 749 ATAAATCTCATTTTCGCGATCTTCCGCGAATCAGTGAAGTGAAGTGAAGTGAAGTGA 808
DB 1070 ATATCTTCTCATTAAGAAATTTTGTGCTGTAAGTGGGCTGACAGTGAAGCAATTTA 1129
QY 809 AGAAATCAATGATTTTTCGATGATTTTGGGGTGGGCGAGAGAGACGACGATTTGG 868
```

```
DB 1130 GAAAGATCAATGTTTTCTTAATGCTTCTGGGAGTGGGAGAGAGATGATGACTTT 1189
QY 869 CGACGAGATCATGATGCTGAGCTGAAGTTTCAAGATATCCGACACAAATTCACGAT 928
DB 1190 GGAACGAGTTCACTATGCTGATGATTAATGAACGACGAGAGAGCTTAAGAAAT 1249
QY 929 ATAAATGATTAAGCACTGACGAGAGAGAGCAATTCAGTTAATTAATCCGCTACAAA 988
DB 1250 ACAAGTCAATCTCTCA-TCACCATAGAGGTGAAGTCAAGTT--TTAGAGCGTATTAAT 1306
QY 989 TAATGGCCCAAGAGAGCGCGATGACACGTCAGCGCTTAAGCATCTGAAGTAAAC 1048
DB 1307 TACTAAGTATTTCCAGAGAGCGTCAATCATGATGATGATGATTAATTAATATGAGC 1366
QY 1049 TCGTAATCTGG 1060
DB 1367 CAAAAATCTGG 1378

RESULT 8
US-09-745-763-84
; Sequence 84, Application US/09745763
; Patent No. US20020065394A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; Lavallee, Edward R.
; Collins-Racie, Lisa A.
; Evans, Cheryl
; Werberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; ENCODING THEM
; NUMBER OF SEQUENCES: 219
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/745,763
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1853 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-09-745-763-84

Query Match 13.6%; Score 157.2; DB 3; Length 1853;
Best Local Similarity 54.4%; Pred. No. 1.3e-39;
Matches 360; Conservative 0; Mismatches 298; Indels 4; Gaps 2;

QY 421 CTTAAGATTTGTTGCAAGGCAATCGTGTGCTATTATTTGCTTATGAGATCGTGA 480
```


Db 1001 AATGAGTGAACGACGACGATGAGCTTTACACCAAGTCACAGCTCTGAGACACA 1060
Qy 1021 GAGCGCTTGAAGCATCTGAAGTCTGTAATCTGGAATTGAAGCTCTTACACT 1080
Db 1061 GATGGTTAGTAGTGTCTTATTAATTAATGATCTGTGACACACATCTTATATATAC 1120
Qy 1081 CGAGCGCTGCTGATTT 1097
Db 1121 AACATCAGATGGATTT 1137

RESULT 10
US-09-989-723-235
; Sequence 235, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guirney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C62
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/087738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948

1	PRIOR FILING DATE: 1998-06-19	60/089952
2	PRIOR APPLICATION NUMBER: 60/089952	
3	PRIOR FILING DATE: 1998-06-19	
4	PRIOR APPLICATION NUMBER: 60/090246	
5	PRIOR FILING DATE: 1998-06-22	
6	PRIOR APPLICATION NUMBER: 60/090252	
7	PRIOR FILING DATE: 1998-06-22	
8	PRIOR APPLICATION NUMBER: 60/090254	
9	PRIOR FILING DATE: 1998-06-22	
10	PRIOR APPLICATION NUMBER: 60/090349	
11	PRIOR FILING DATE: 1998-06-23	
12	PRIOR APPLICATION NUMBER: 60/090355	
13	PRIOR FILING DATE: 1998-06-23	
14	PRIOR APPLICATION NUMBER: 60/090429	
15	PRIOR FILING DATE: 1998-06-24	
16	PRIOR APPLICATION NUMBER: 60/090431	
17	PRIOR FILING DATE: 1998-06-24	
18	PRIOR APPLICATION NUMBER: 60/090444	
19	PRIOR FILING DATE: 1998-06-24	
20	PRIOR APPLICATION NUMBER: 60/090445	
21	PRIOR FILING DATE: 1998-06-24	
22	PRIOR APPLICATION NUMBER: 60/090472	
23	PRIOR FILING DATE: 1998-06-24	
24	PRIOR APPLICATION NUMBER: 60/090535	
25	PRIOR FILING DATE: 1998-06-24	
26	PRIOR APPLICATION NUMBER: 60/090540	
27	PRIOR FILING DATE: 1998-06-24	
28	PRIOR APPLICATION NUMBER: 60/090542	
29	PRIOR FILING DATE: 1998-06-24	
30	PRIOR APPLICATION NUMBER: 60/090557	
31	PRIOR FILING DATE: 1998-06-24	
32	PRIOR APPLICATION NUMBER: 60/090676	
33	PRIOR FILING DATE: 1998-06-25	
34	PRIOR APPLICATION NUMBER: 60/090678	
35	PRIOR FILING DATE: 1998-06-25	
36	PRIOR APPLICATION NUMBER: 60/090690	
37	PRIOR FILING DATE: 1998-06-25	
38	PRIOR APPLICATION NUMBER: 60/090694	
39	PRIOR FILING DATE: 1998-06-25	
40	PRIOR APPLICATION NUMBER: 60/090695	
41	PRIOR FILING DATE: 1998-06-25	
42	PRIOR APPLICATION NUMBER: 60/090696	
43	PRIOR FILING DATE: 1998-06-25	
44	PRIOR APPLICATION NUMBER: 60/090682	
45	PRIOR FILING DATE: 1998-06-26	
46	PRIOR APPLICATION NUMBER: 60/090863	
47	PRIOR FILING DATE: 1998-06-26	
48	PRIOR APPLICATION NUMBER: 60/091360	
49	PRIOR FILING DATE: 1998-07-01	
50	PRIOR APPLICATION NUMBER: 60/091478	
51	PRIOR FILING DATE: 1998-07-02	
52	PRIOR APPLICATION NUMBER: 60/091544	
53	PRIOR FILING DATE: 1998-07-01	
54	PRIOR APPLICATION NUMBER: 60/091519	
55	PRIOR FILING DATE: 1998-07-02	
56	PRIOR APPLICATION NUMBER: 60/091626	
57	PRIOR FILING DATE: 1998-07-02	
58	PRIOR APPLICATION NUMBER: 60/091633	
59	PRIOR FILING DATE: 1998-07-02	
60	PRIOR APPLICATION NUMBER: 60/091787	
61	PRIOR FILING DATE: 1998-07-07	
62	PRIOR APPLICATION NUMBER: 60/091982	
63	PRIOR FILING DATE: 1998-07-07	
64	PRIOR APPLICATION NUMBER: 60/092162	
65	PRIOR FILING DATE: 1998-07-09	

	Query Match	13.5%	Score 155.4	DB 3	Length 1964
	Best Local Similarity	51.8%	Pred. No. 5.3e-39		
	Matches 351, Conservative	0;	Mismatches 326;	Indels 0;	Gaps 0
OY	421 COTAGAGTGTGTTGCAGGCGATCGTGTCTATTATTATTCGCCCTTAGAGACGTGA	480			

Db	461	CGCTAGGAATGTAAGCTTTACAGAGGGTCGCCATCCTCGTTCCCA	CCGGAACAGAGAG	520
Qy	481	GCACATTTGAGATTAATGCTTCACAAATTTGCACTCGTTGCTCGCCAAACAATTTGAC		540
Db	521	AAACACCTGATGTAACCTGCTGGAACATCTGCATCCTTCCTCGCAGAGCGACACTGGAT		580
Qy	541	TATGCATTTTCAATGTCGAGCAAGTGGCGCAATCAGACGTTTAATCGCGGAAACTAATG		600
Db	581	TATGGCATCTACCTGCATCCACACAGGCTGAAGGTAAAAAGTTTAATCGACCAAACTCTTG		640
Qy	601	AACCTGTGATACGACGTAGCATCAAGCCTCTACCCATGGCAGTGCCTTCATCTTCATGAT		660
Db	641	AATGTGGGCTATCTAAGAGCCCTCAAGGAAGAAATATGGGACTGCTTATATTCACAGAT		700
Qy	661	GTCGATTTACTGCCCGGAAGATGACCGTAACCTGTACACGTGTCCAAATTCACACACGTAT		720
Db	701	GTGGAACCTGTGATCCCGGAATGACTTTTAACCTTTACAAGTGTGAGAGCATCCCAAGCAT		760
Qy	721	ATGAGTGTAGGATCGCATTAATTCATTAATTAATTCATTAATTTGGGGATCTTGGCGGA		780
Db	761	CTGGTGTGTGGCAAGAACAGCATGGGTGGTACGGTTACGTTCAGGTGATATTTTGGGGGT		820
Qy	781	ATCAGTGCACATCAAAAGATTCACCTGAAGAAATCAATGGAATTTGCAATGATTTTGG		840
Db	821	GTTACTGCCCTACACAGAGCACTTTTTCAGAGTGAATGSAATTTCTTAACAACACTACGG		880
Qy	841	GATTGGGCGGAGAGACGACGATTTGGCGACGAAATCATGATGGCTGGACTGAAGTT		900
Db	881	GGATGGGGAGGCGAAGCGATGACCTGACCTCAGGCTGAGGCTCCAAAGATGAAAATT		940
Qy	901	TCAAGATATCCGACACAAATTCGACGATTAATTAATGATTAAGCATCGTAGGGAAAGCGAG		960
Db	941	TCCCGGCCCTGCTGAAGTGGGTAAATATACAAATGGTCTTCCACACTAGAACAAAGGC		1000
Qy	961	AATCCAGTTATAATATGCCGCTACAAAATAATGSGCCAAACGAAGCGCCGATGACAGT		1020
Db	1001	AATAGAGTGAACGAGAACGAGATGAAGCTCTTACACCAAGTGTACGAGGTGTGGAGAAC		1060
Qy	1021	GACGGCTTAAGCAATCTGAAGTATPAAGCTGTAAATCTGGAATTTGAAGCCTCTCTACACT		1080
Db	1061	GATGGGTGAATGATGTTGTTTATATAATATGATATCTGTGAACACAAATCTTATATATATC		1120
Qy	1081	CGAGCCGCTGTCGATTT	1097	
Db	1121	AACATCAACAGTGAATTT	1137	

```

RESULT 11
US-09-989-279-235
Sequence 235, Application US/09989277235
Patent No. US20020072496A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Boetstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Andrew
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J Christopher
APPLICANT: Gurney, Auetin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel

```

APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PLC56
CURRENT APPLICATION NUMBER: US/09/989,279
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678

```

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090693
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

Query Match 13.5%; Score 155.4; DB 3; Length 1964;

Best Local Similarity 51.8%; Pred. No. 5.3e-19; Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

```

Qy 421 CCTAAGATTGTGTGCAAGCATCGTGTCTATTAATGTGCTTATAGATCGTGA 480
Db 461 CCTCAGGAATGTAAAGTTTACAGAGGGTCCCATCTCTCCCGCAAGAGAG 520
Qy 481 GCACATTTGGAATTAATGCTTCCCAATTTGCACTCGTGTCCGCAACAATTTGAC 540
Db 521 AACACCTGTATGTACCTGCGAACAATCTGATCCCTTCTGCAAGGCGAGCTGGAT 580
Qy 541 TATGCAATTTCTATGTGAGCAAGTGGCGAATCGAGCTTTATCGGGAAATCTATG 600
Db 581 TATGCAATTTCTATGTGAGCAAGTGGCGAATCGAGCTTTATCGGGAAATCTATG 640
Qy 601 AACGTTGATATGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 641 AATGTGGGCTATCTAGAAAGCCCTCAAGAAAGAAATTTGGGAGCTTTATATTCAGAT 700
Qy 661 GTGATTTATGCTGCGGAAAGTGAACCGTACCTGTGACGCTGTGCAATTCACGTCAT 720
Db 701 GTGACCTGTATGCTGCGGAAAGTGAACCGTACCTGTGCAATTCACGTCAT 760
Qy 721 ATGAGTTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 761 CTGTGTGTTGAGCAAGACAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Qy 781 ATGAGTTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 821 GTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
Qy 841 GGTGGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 881 GGTGGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
Qy 901 TCAAGATATCGACACAAATTTGACGATPATAATGATTAAGCTGAGAGAGAGAG 960
Db 941 TCCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Qy 961 AATCAGTTAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

```

```

Db 1001 AATGAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1060
Qy 1021 GAGGCGCTTACGAATCTGAGATTAAGCTCGTAATCTGGAATTAAGCTCTACACT 1080
Db 1061 GATGGGTGAGTATGTTGTTCTTATTAATGATGATGATGATGATGATGATGATGAT 1120
Qy 1081 CGAGCGCTGCTGATTT 1097
Db 1121 AACATCAGTGTGATTT 1137

```

RESULT 12

US-09-989-727-235

Sequence 235, Application US/09989727

Patent No. US2002072497A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerltzen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Steward, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2730P1C65

CURRENT APPLICATION NUMBER: US/09/989,727

CURRENT FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19

PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090676
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090678
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090690
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090696
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090862
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
 Best Local Similarity 51.8%; Pred: No. 5.3e-39;
 Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTGTCGAAGCATCGTGTGCTATTATGTCCTTAAGATCGTGA 480
DB 461 CCTCAGGAATGTAAAGCTTACAGAGGTCGTCATCGTGTCCCAACCGAACAAGAG 520
QY 481 GCACTATTGAGATTAATGCTCCCAATTGGCACTCGTGTGCCCAACAACATTGGAC 540
DB 521 AAACACCTTAATGACCTGCGAACAATCGCACTCCCTTCTGAGAGGCGAGCGTGGAT 580
QY 541 TATGCAATTTCATTGTGAGCAAGTGGCAATCAGCGTTTATGCGGGAAATATAG 600
DB 581 TATGCACTTACGTCATCCACGAGCTGAAGTAAAGTTTATGAGCAAACTCTTG 640
QY 601 AACGTTGATACAGCAGTACGATCAGCGCTTACCAATGCGAGCTTCACTTTCATGAT 660
DB 641 AATGTGGCTATCTAGAGGCCCTCAAGAAAGAAATTTGGAGCTTTATATCCAGAT 700
QY 661 GTGCAATTACGCCCCGAAGTGAACCGTAACCTGTACACGTCCTCAATTCACACGTCAT 720
DB 701 GTGAGACCTGTACCCGGAATGACTTTACCTTTACAAAGTGAAGAGAGCAATCCAGCAT 760
QY 721 ATGAGTGTAGCGATCGATTAATTAATTCAATTATTAATTCATATTCGCGCATCTTGGCGGA 780
DB 761 CTGTGTGTGTCAGAGAACGACACTGGGTACGTTACGTTACGTTATTTTGGGGGT 820
QY 781 ATGAGTGCATTAACAAAGATCACTGAAGAAAATCAATGATTTTGAATTTTGG 840
DB 821 GTTACGTCCCTTAAGCAGAGAGCGATTTTCAAGGTGAATGGAATCTCTTAACAACACTGG 880
QY 841 GGTGGGGGGCGAAGACGACGATTTGGCGAACGAATGCGTGAACGTAAGTT 900
DB 881 GGATGGGGAAGCGAAGACGATGACCTGACGAGGTTGAGCTCCAAAGAAATTT 940
QY 901 TCAAGATATCCGACACAAATTTGACGATATAATGATTAAGACTTCGACGAGGACG 960
DB 941 TCCCGGCCCCCTGCTGAATGGGTAATATACATGCTCTTCAACATAAGCAAAAGC 1000
QY 961 AATCCAGTTAATAATGCGCTCAAAAATAATGGGCCAAAGAACGCGCATGACGTC 1020
DB 1001 AATGAGGTGAACGACGAAGAGCTCTTACACCAAGTGCACGAGTCTGGAACA 1060
QY 1021 GAGGCGCTAAGCAATCGAAGTAACTGTAAATTTGAAATTTGAGCTCTCTACAT 1080
DB 1061 GATGGGTGAGTGTGTTCTTTAATAATGATCTGTGAACAACATCTTTATATATC 1120
QY 1081 CGAGCGCGTCGATTT 1097
DB 1121 AACATCACAGTGAATTT 1137

RESULT 13

US-09-989-731-235
Sequence 235, Application us/09989731
Patent No. US20020103125A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PLC70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090676
 PRIOR FILING DATE: 1998-06-25

PRIOR APPLICATION NUMBER: 60/090678
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090690
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090694
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090695
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090696
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090862
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
 Best Local Similarity 51.8%; Pred. No. 5.3e-39;
 Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

421 CCTAAGATTGTTGTTGCAAGGCAATGCTGCTATTTATTTGSCCTATAGAGATCGTGA 480
 461 CCTCAGATGTAAGGCTTTACAGAGGTCGCTATCCTCCCAACGGAACAGAG 520
 481 GCACATTTGAGAAATATGCTCCCAATTTGCACTGCTGCTGCCAACAACATTTGAC 540
 521 AAACACCTGATGTACCTGTGGAACATCGATCCCTTCTCAGAGGAGCAGCGTGGAT 580
 541 TATGCAATTTTATTGTTGAGCAAGTGGCAATCGACGTTTAATCGCGGAACTAATG 600
 581 TATGCAATCTACGTATCCACCAAGCTGAAGGTAAAGTTAAATCGAACCAACTCTTG 640
 601 AACGTTGATTAAGAGTATGATCAAGCGCTCTACCGATGCGATGCTTTCATGAT 660
 641 AATGCGGCTATCTGTAAGAGCCCTCAAGAGAAATTTGGATCTGCTTATTTCCAGAT 700
 661 GTGATTTTACTGCGCGAAGATGACGTAACCTGTACAGGTGTCAATTCACACGTCAT 720
 701 GTGACCTGTTACCGAGATGACTTTAACTTTCAAGGTGAGAGATCCCAAGAT 760
 721 ATGAGTATGCGATGATTAATTTCAATTTAACTTTCAATTTGCGGATCTTGGCGGA 780
 761 CTGCTGTTGTTGCGAGAACGCACTGGTACAGTTACGTAAGGATATTTGGGGGT 820
 781 ATCAGTGACATTAACAAAGATCACTGAAGAAATCAATGATTTTTCGATGATTTTGG 840
 821 GTTATCTGCTTACAGAGAGAGAGATTTTCAAGGTGATTTCTTAACTAATCTG 880
 841 GGTGGGGGCGAGAGAGAGAGATTTTGGCGAGAGAACTGATGATGAAAGTT 900
 881 GGATGGGGGAGAGAGAGAGATGACTCAAGCTCAGGTTGAGCTCAAGATGAAATTT 940
 901 TTAAGATATCCAGCAAAATTTGACGATTTAAATGATTAAGCATCTGACGGAAGCAG 960
 941 TCCTGGCCCTGCTGTAAGGTGTTAAATTAATGATGTTTTCACATGAGAGCAAAAGC 1000

PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 5.3e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATGTTGTTGCAAGGCGATGTTGCTATTAATGTCCTTAAGAGTCGTAA 480
DB 461 CCTCAGAGATGTTAAAGCTTTACAGAGGTCGCTCCTTCCCAAGGAGAGAG 520
QY 481 GCACATTTGAGATTAATGTCACAAATTTGCACTGTTGTCGCGCAACAAATTTGAC 540
DB 521 AAACACCTGATGTATCTGCTGGAACATCTGATCCCTTCTCAGAGGAGAGCTGAT 580
QY 541 TATGCAATTTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 581 TATGCAATTTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
QY 601 AACGTTGATTAAG 660
DB 641 AATGTCGCTATCTGAG 700
QY 661 GTGATTTATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 701 GTGACCTGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
QY 721 ATGATGTTAGAGATGATTAATTTCAATTTAATTTCAATTTCAATTTCAATTTCA 780
DB 761 CTGATGTTGAG 820
QY 781 ATCAGTGAATTAAG 840
DB 821 GTTATCTGCTTACAG 880
QY 841 GGTGAG 900
DB 881 GATGAG 940
QY 901 TCAAGATTTGAG 960
DB 941 TCCGAG 1000
QY 961 AATCAGTTAATTAAG 1020
DB 1001 AATGAGTTGAG 1060
QY 1021 GACGAGCTTAAG 1080
DB 1061 GATGAGTTGAG 1120
QY 1081 CGAGCTGCTGATTT 1097
DB 1121 AACATCAG 1137

RESULT 15
US-09-991-073-235
Sequence 235, Application US/09991073
Patent No. US20020127576A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ealon, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavits, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C15
CURRENT APPLICATION NUMBER: US/09/1991, 073
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 18:50:05 : Search time 243 Seconds
(without alignments)
8426.959 Million cell updates/sec

Title: US-10-661-430-2

Perfect score: 1152

Sequence: 1 atgcttctcgtcattcgtgc.....acttccacgctgttttag 1152

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6/COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6/COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6/COMB.seq:*
6: /cgn2_6/ptodata/1/ina/6/COMB.seq:*
7: /cgn2_6/ptodata/1/ina/6/COMB.seq:*
8: /cgn2_6/ptodata/1/ina/6/COMB.seq:*
9: /cgn2_6/ptodata/1/ina/6/COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	16.5	1189	3	US-09-270-767-11472
2	167.6	14.5	1938	3	US-09-673-395A-140
3	167.6	14.5	1970	3	US-09-949-016-2875
4	161.2	13.0	3830	3	US-09-949-016-679
5	159.6	13.9	3832	3	US-09-949-016-2359
6	155.4	13.5	1964	3	US-09-991-181-235
7	155.4	13.5	1964	3	US-09-990-444-235
8	155.4	13.5	1964	3	US-09-997-333-235
9	155.4	13.5	1964	3	US-09-992-598-235
10	155.4	13.5	2151	3	US-09-949-016-1167
11	155.4	13.5	2167	3	US-09-949-016-541
12	155.4	13.5	2280	2	US-09-055-097-4
13	155.4	13.5	2280	2	US-09-373-903-4
14	150	13.0	2265	3	US-09-949-016-4375
15	150	13.0	4646	3	US-09-949-016-680
16	139.6	12.1	1265	2	US-08-446-777-1
17	139.6	12.1	2304	2	US-08-446-777-5
18	139.6	12.1	2304	2	US-08-446-777-7
19	134	11.6	1023	3	US-09-118-464-34
20	134	11.6	1023	3	US-10-132-652A-34
21	134	11.6	1116	3	US-09-118-464-33
22	134	11.6	1116	3	US-10-132-652A-33
23	134	11.6	1119	3	US-09-949-016-2195
24	134	11.6	1119	3	US-09-949-016-2195

25	134	11.6	2027	3	US-09-118-464-1	Sequence 1, Appl1
26	134	11.6	2027	3	US-10-132-652A-1	Sequence 1, Appl1
27	131.6	11.4	494	3	US-09-270-767-27061	Sequence 27061, A
28	108.8	9.4	541	3	US-09-404-879A-108	Sequence 108, App
29	108.8	9.4	541	3	US-09-338-933-108	Sequence 108, App
30	108.8	9.4	541	3	US-09-215-681-108	Sequence 108, App
31	108.8	9.4	541	3	US-09-215-681-108	Sequence 108, App
32	108.8	9.4	541	3	US-09-215-681-108	Sequence 108, App
33	108.8	9.4	541	3	US-10-198-053-108	Sequence 108, App
34	108.8	9.4	541	3	US-09-827-271-108	Sequence 108, App
35	69.4	6.0	500	3	US-09-270-767-1933	Sequence 1933, Ap
36	69.4	6.0	500	3	US-09-270-767-1933	Sequence 1933, Ap
37	66.8	5.8	26313	3	US-09-949-016-16117	Sequence 16117, A
38	66.8	5.8	81384	3	US-09-949-016-12422	Sequence 12422, A
39	58	5.0	601	3	US-09-949-016-81352	Sequence 81352, A
40	58	5.0	601	3	US-09-949-016-81352	Sequence 81352, A
41	58	5.0	65518	3	US-09-949-016-14101	Sequence 14101, A
42	57.2	5.0	601	3	US-09-949-016-31230	Sequence 31230, A
43	57.2	5.0	601	3	US-09-949-016-31231	Sequence 31231, A
44	57.2	5.0	601	3	US-09-949-016-81354	Sequence 81354, A
45	57.2	5.0	601	3	US-09-949-016-81354	Sequence 81354, A

ALIGNMENTS

RESULT 1									
US-09-270-767-11472									
; Sequence 11472, Application US/09270767									
; Patent No. 6703491									
; GENERAL INFORMATION:									
; APPLICANT: Homburger et al.									
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster									
; FILE REFERENCE: File Reference: 7326-094									
; CURRENT FILING DATE: 1999-03-17									
; NUMBER OF SEQ ID NOS: 62517									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 11472									
; LENGTH: 1189									
; TYPE: DNA									
; ORGANISM: Drosophila melanogaster									
US-09-270-767-11472									
Query Match 16.6%; Score 191; DB 3; Length 1189;									
Best Local Similarity 58.6%; Pred. No. 1.2e-51;									
Matches 332; Conservative 0; Mismatches 235; Indels 0; Gaps 0;									
QY	399	GCACGCCGATGACATGGAATGCTTAAGATTTGTTGCAAGCAGTGGTTGATTAT	458						
DB	508	GCACGCCGATGACATGGAATGCTTAAGATTTGTTGCAAGCAGTGGTTGATTAT	567						
QY	459	TGTGCTTATGAGATGCTGGAAGCATTTGGAATTAATGCTCAAAATTTGACTGTT	518						
DB	568	TGTGCTTATGAGATGCTGGAAGCATTTGGAATTAATGCTCAAAATTTGACTGTT	627						
QY	519	GCTGCCCAACCAATGATGATGATTTTATTTGAGAGCAAGTGGCAATCAAGC	578						
DB	628	TCTGATGAGAGCAAGTGGATGCTTATGATTTTATTTGAGAGCAAGTGGCAATCAAGC	687						
QY	579	GTTTATGCGCGGAATGAACTGATGATGATGATGATGATGATGATGATGATGATG	638						
DB	688	CTTCAATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	747						
QY	639	GCAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	698						
DB	748	GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	807						
QY	699	GTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	758						
DB	808	CTGTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	867						
QY	759	ATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	818						

Db 868 TTATCATCAATATTTGGAGGTGTTTCCGCAATGACCGCGAGACATTTCCAGCCCGTA 927
Qy 819 TGGATTTTGAATGATTTTGGGGTGGCGGAGAGAGAGAGATTTGGCCAGCAAGAAC 878
Db 928 TGGGTTCTAATACTGTTCTTTGGCTGGGGCGGCGAGAGATGAGACATGTCCTCAACGGTT 987
Qy 879 ATCGATGGCTGAGCTGAAGTTTCAAGATATCCGACACAAATTGACGATATAAATGAT 938
Db 988 GAAGACGCCCACTTATTCATCAAGGATCCGGTCAACATAGCCCGCTCAAGATGCT 1047
Qy 939 TAAGCACTGACGGAAGCAAGATCC 965
Db 1048 GAAGCATCAGAAAGAAAGGCAATCC 1074

RESULT 2

US-09-673-395A-140
; Sequence 140, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BRAND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673, 395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 140
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-673-395A-140

Query Match 14.5%; Score 167.6; DB 3; Length 1938;

Best Local Similarity 54.4%; Pred. No. 7.1e-44; Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

Qy 421 CCTAAGGATTGTGTGCAAGGACATCGTGTCTATTATTGTGCGCTATAGAGATCGTGA 480
Db 573 CTCGAGGTTGTAGCCCCCGCTCCGCAACAGCATATTGTGCTCATCGTCCCGGAG 632
Qy 481 GCACATTGAGAAATATGCTTCCACAAATTTGCACTGTTGCTGCCAAACAAATGGAC 540
Db 633 CACCACTGCGCTGCTCTCAACCTGACACCTTCTTGGACGCGCAGCACTTGTCT 692
Qy 541 TATGCAATTTTCAATGTGGAAGCAAGTGGCAATGAGAGTTTATGCGGGAACTAATG 600
Db 693 TATGGAATCTATGATCATCCACAGGCTGGAATGAACATTTAACAGGGCAAACTTTG 752
Qy 601 AACGTTGATAGCACTGATGATCAAGCTCTTACCCATGAGAGTGTTCATCTTTCATGAT 660
Db 753 AAGTTGGGTGGAGAGGCGCTGCGATGAAGAAGTGGAGCTGCTCTTTCGACGAT 812
Qy 661 GTGCAATTAAGTCCGGAAGATGACGTTAAGCTGTACAGTGT---CCAATTCAACAGT 717
Db 813 GTGGAATCTTGGCAGAAATGACCAATCTGTATGTGTGACCCCGGGGACCCGCG 872
Qy 718 CATATGAGTGAAGGATCGAATTAATGATTAATTAATTCATTCATTTGGGAGTCTGGC 777
Db 873 CATGTGCGCTGTGTATGAACAAAGTTGATGACGCTCCCGTACCCCGAGTACTTGA 932
Qy 778 GAATCAGTGAAGCAAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTT 837
Db 933 GGAATCTCAGACATTAATCTTCCGACCAATGATGAAATGATGCTTCCCAATGATAC 992
Qy 838 TGGGTTGGGGCGGAGAGACGATTTGGCGACGAAACATGATGCTGACGTGAA 897

Db 993 TGGGCTGGGGTGTGAGAGATGACGATTTGCTACCCAGGTCGCCGTGGATGAAG 1052
Qy 898 GTTTCAGATATCCGACCAAAATGACGATATAAATGATTAAGCATCGACGAGGCG 957
Db 1053 ATCTTCGGCCCCCACAATCTGTAGACACTTAATGATGTGAAGCAGCAGAGATGAAG 1112
Qy 958 ACGAATCCAGTTAATTAATGCGCTACAAATATATGGGCAAAAGAGCCGATGACA 1017
Db 1113 GGCATAGAGAAATATCCACAGATTTGACCTCTGTGTCGTAACCAATTCCTGACG 1172
Qy 1018 CGTAGCGCTTAAGCAATGTGAATGATCTGTAAATCTGGAATGAAGCTCTCTAC 1077
Db 1173 CAAGATGGAGTAACTCACTGACATACAGTTGTGGCTGAGAGCTGGGCTCTTAT 1232
Qy 1078 AC 1079
Db 1233 AC 1234

RESULT 3

US-09-949-016-2875
; Sequence 2875, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2875
; LENGTH: 1970
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2875

Query Match 14.5%; Score 167.6; DB 3; Length 1970;

Best Local Similarity 54.4%; Pred. No. 7.2e-44; Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

Qy 421 CCTAAGGATTGTGTGCAAGGACATCGTGTCTATTATTGTGCGCTATAGAGATCGTGA 480
Db 956 CTCGAGGTTGTAGCCCCCGCTCCGCAACAGCATATTGTGCTCATCGTCCCGGAG 1015
Qy 481 GCACATTGAGAAATATGCTTCCCAATTTGCACTGTTGCTGCCAAACAAATGGAC 540
Db 1016 CACCACTGCGCTGCTCTCAACCTGACACCTTCTTGGACGCGCAGCACTTGTCT 1075
Qy 541 TATGCAATTTTCAATGTGGAAGCAAGTGGCAATGAGAGTTTATGCGGGAACTAATG 600
Db 1076 TATGGAATCTATGATCATCCACAGGCTGGAATGAACATTTAACAGGGCAAACTTTG 1135
Qy 601 AACGTTGATAGCACTGATGATCAAGCTCTTACCCATGAGAGTGTTCATCTTTCATGAT 660
Db 1136 AACGTTGGGTGGAGAGGCGCTGCGATGAAGAAGTGGAGCTGCTTTCGACGAT 1195
Qy 661 GTGCAATTAAGTCCGGAAGATGACGTTAAGCTGTACAGTGT---CCAATTCAACAGT 717
Db 1196 GTGGAATCTTGGCAGAAATGACCAATCTGTATGTGTGACCCCGGGGACCCGCG 1255
Qy 718 CATATGAGTGAAGCAATGATTAATTAATTAATTAATTCATATTTGGGAGATCTTGGC 777
Db 1256 CATGTGCGCTGTGTATGAACAAAGTTGATGACGCTCCCGTACCCCGAGTACTTGA 1315
Qy 778 GAATCAGTGAAGCAAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTT 837

DB 1316 GGAAGTCTAGGACTACTCTGACCAAGTACCTGAAGATGATGGCTTCCCAATGAAATAC 1375
QY 838 TGGGGTTGGGGGAGAGAGACAGATTTGGGACAGAGAACTGCATGGCTTGACCTGAA 897
DB 1376 TGGGGCTGGGGTGGTGGATGATGACGATTTGCTACAGGGTGGCGCTGGCTGGATGAAG 1435
QY 898 GTTTCAGATATCCGACAAATGTCACGATATTAATGATTAAGCACTGACGAGAGCG 957
DB 1436 ATCTCTGGCCCCCAGATCTGTAGGACCTATTAAGTGTGAAGCAGGAGAGATTAAG 1495
QY 958 ACGAATCCAGTTAATTAATGCGCTACAAATTAATGGGCGCAAGAGCGCGATGACA 1017
DB 1496 GGCATGAGAGAAATCCCAAGATTTGACCTCTGCTCCGTACCCAGAAATTCCTGAGAG 1555
QY 1018 CCGTACGCGCTTACGACATTTGAAGTAACTGCTTAATCTGGAATGAAAGCTCTTAC 1077
DB 1556 CAAGATGGAGTAACTGACATGACATGACGTTGCTGAGAGCTGGGGCTCTTAT 1615
QY 1078 AC 1079
DB 1616 AC 1617

RESULT 4
US-09-949-016-679
; Sequence 679, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 679
; LENGTH: 3830
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-679

Query Match 14.0%; Score 161.2; DB 3; Length 3830;
Best Local Similarity 54.5%; Pred. No. 1.3e-41;
Matches 366; Conservative 0; Mismatches 303; Indels 3; Gaps 2;
QY 389 ATCCGACACGACCGCGGTGACATGGAATGCTTAAGATTTGTTGCAAGGCATCGTG 448
DB 652 ATTGATATTTAGAGCAGGGGCTCATTTGAGGCCAAAGACGTAAACCAATGGAAG 711
QY 449 TTGCTATTATTTGCTTATGAGATCGTGAAGCAGATTTGGAATATGCTCCCAATT 508
DB 712 TGGCAGTTCTCATTTCTTCCGTAATCGCCATGAACATCTCCAAATTTTCTTACATC 771
QY 509 TGCACCTGTTGCTCGCAAAACAATTGAGCTATGCAATTTTCAATGTGAGCAAGTGG 568
DB 772 TGATTTCAATGCTCCAGAGAGACGCGCTGGAATTTGCGTTTATGTGTTGAACAGACTG 831
QY 569 CGAATCAGAGCTTTAATCGCGGAAACTAATGACGTTGATACGACGATGACGCGC 628
DB 832 GCAACAACTTTTAAACGTTGAGTCTTTCAATGTGGCTTCAAAGAGGCGCATGAAG 891
QY 629 TCTACCATGAGCAGTCTTCAATGATGATGATGATTTTACCTGCCGAAGTACCGTA 688
DB 892 ACAGTGTCTGGAGCTGTGTATCTTCCAGATGTGATCATCTACCTGAATAATGACGGA 951

QY 689 ACTGTACAGCTGTCCAAATTCACCAAGCTCATATGATGATGACATCAATTAATCAATT 748
DB 952 ACTATTAGAGATGTGAGAAATGCGACGTCAATTTTGTGCAAGCTGATTAATACATGT 1011
QY 749 ATAACTTTCATATTGCGCGATCTTGGCGGATATGACACTTAACAAAGATCACTGA 808
DB 1012 ATATCTTCCATATTAATAAATTTTGTGTGTATGAGGGCTGACAGTGAACAATTTA 1071
QY 809 AGAAATCAATGAGATTTTGCATGATTTTGGGGGTTGGGGGAGAGGCGACGATTTGG 868
DB 1072 GAAAGATCAATGATTTTCTTAATGCTTCTGTGGAGATGGGAGAGAAATGATGACTTT 1131
QY 869 CGACGAGAACTCATGATGCTGACCTGAAGATTTCAAGATATCCGACACAAATTTGACGAT 928
DB 1132 GGAACAGATTCACATCTCTGATTAATGTAACGACGAGGAGGAGCTTAGAATAAT 1191
QY 929 ATAAATATTTAGACATCTGACGAGAGGAGGAGATTCGATTAATATGCGCTCAAAA 988
DB 1192 ACGATCAATTCCTCA-TCACCATGAGGTGAAGTCCAGTT--TTTAGACGCTATTAAT 1248
QY 989 TAATGGGCGCAAGAGGCGGATGACACGTCAGCGCTAAGCAATCTGAATTAAGC 1048
DB 1249 TACTAAGGTATTCAGAGAGGCTCAGTACATGATGATGACGACATTAATATATAGGC 1308
QY 1049 TCGTAATCTGG 1060
DB 1309 CAAAATACTGG 1320

RESULT 5
US-09-949-016-2359
; Sequence 2359, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2359
; LENGTH: 3832
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2359

Query Match 13.9%; Score 159.6; DB 3; Length 3832;
Best Local Similarity 54.3%; Pred. No. 4.2e-41;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;
QY 389 ATCCGACACGACCGCGGTGACATGGAATGCTTAAGATTTGTTGCAAGGCATCGTG 448
DB 652 ATTGATATTTAGAGCAGGGGCTCATTTGAGGCCAAAGACGTAAACCAATGGAAG 711
QY 449 TTGCTATTATTTGCTTATGAGATCGTGAAGCAGATTTGGAATATGCTCCCAATT 508
DB 712 TGGCAGTTCTCATTTCTTCCGTAATCGCCATGAACATCTCCAAATTTTCTTACATC 771
QY 509 TGCACCTGTTGCTCGCAAAACAATTGAGCTATGCAATTTTCAATGTGAGCAAGTGG 568
DB 772 TGATTTCAATGCTCCAGAGAGACGCGCTGGAATTTGCGTTTATGTCAATGAAACAGACTG 831
QY 569 CGAATCAGAGCTTTAATCGCGGAAACTAATGACGTTGATACGATGACGATGACGCGC 628
DB 832 GCAACAACTTTTAAACGTTGAGTCTTTCAATGTGGGCTTCAAAGAGGCGCATGAAG 891

[illegible]

PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Query Match	13.5%	Score 155.4	DB 3	length 1964
Best Local Similarity	51.8%	Pred. No. 7e-40		
Matches 351, Conservative	0	Mismatches 326	Indels 0	Gaps 0

Qy	421	CTTAAGGATTTGTGTGCAAGGCATTCGGTTGCTATTATGTGACCTTAAAGATCGTAA	480
Db	461	CTTCAGGAATGTAAAGCTTTACAGAGGGTGCATCTCTGTTCCTCACCGAAACAGAG	520
Qy	481	GCACTTTGGAATATAGTCTCCACAAATTTGCACGTGTCTCGCCAAAACAATTGAC	540
Db	521	AAACACTGTATGTACCTGCTGGAACATCTGCATCTCTCTGCGAAGGACAGCACTGGAT	580
Qy	541	TATGCAATTTTCAATGTGAGGACAGTGGCCGAATAGAGTTTATTCGGGGAACTAATG	600
Db	581	TATGGCATCTACCTGCATCCACAGGCTGAAGGTAAAAATTATTCAGCCAACTCTTG	640
Qy	601	AACGTGGATACGACGTAGCATCAAGCCTCTACCCATGGCAGTGCCTTATCTTTCATGAT	660
Db	641	AAATGTGGCTATCTAGAAAGCCCTCAAGAGAAATAATGGGACTGCTTTATATTCCAGAT	700
Qy	661	GTGCATTTACTGCCCGGAGATGACCGTAACTCTGACAGTGTCCAAATTCACCAAGTAT	720
Db	701	GTGACCTGTGTACCCGAGAAATGACTTTTAACTTTTACAAGTGTAGAGACATCCCAAGAT	760
Qy	721	ATGAGTGTACGCATTCGATTAATTCATTAATTAATCTTCATATTGGGCAATTTGGCGGA	780
Db	761	CTGGTGTGTGGCAGGAACAGCATGGGTACAGGTACAGTTACATGGATATTTTGGGGGT	820
Qy	781	ATCAGTGCACCTAACCAAAAGATCACTGGAAGAAAAATCATGTGATTTTGCATGATTTTGG	840
Db	821	GTTACTCTCCTTAAGCAGAGAGACAGTTTTCAGGTGATGGATCTCTTAACAACACTACGG	880
Qy	841	GGTTTGGGGGCGAGAGCGACGATTTGGCCGACGGAACATCGATGGCTGCGTAAGTT	900
Db	881	GGATGGGGAAGCCGAGACGATACCTTAGACTCAGGGTTGAGCTCCAAAGATGAAATTT	940
Qy	901	TCAAGATATCCGACACAATTTGACGATATATAATGAATTAGCACTGCACGGAACGACG	960
Db	941	TCCCGGCCCTGCTGAAGTGGTAATATATACATGTGCTTCCACACTAGGACAAAGGC	1000
Qy	961	AATTCAGTTAATAATGTCGCTACAAAATAATGGGCCAAACGAAGCGCCGATGGAACGT	1020
Db	1001	AATGAGGTGAACGACGAACGAGATGAACCTCTTACACCAAGTGCACGAGTCTGGAAGAAC	1060
Qy	1021	GACGGCCTTAAGCAATCGAAGTATTAAGCTGTAAATCTGAAATGGAAGCTCTTCACT	1080
Db	1061	GATGGGTGTAGTAGTGTCTTATTAATTTAGTATCTGTGGAACACAAATCTTTATATATC	1120
Qy	1081	CGAGCGGTGTCGATTT 1097	
Db	1121	AACATCAAGTGAATTT 1137	

RESULT 7
US-09-990-444-235

Sequence 235, Application US/09990444
Patent No. 6930170
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C19
CURRENT APPLICATION NUMBER: US/09/990,444
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;

Best Local Similarity 51.8%; Pred. No. 76-40; Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTGTTCAAGGATCGTGTATATATGTCCTATAGAGATCTGAA 480
DB 461 CCTCAGAGATGTAAAGCTTTACAGAGAGGTGCGCATCTGTTCCACCGGAACAGAG 520
QY 481 GCACATTGAGATATATGTCCTCAATTGTCATCTGTCGCCCAACAACATTGAC 540
DB 521 AAACACCTTAGTACCTGCTGGAACATCTGCATCCCTCTGAGAGGAGCAGCTGAT 580
QY 541 TATGCAATTTTCAATGTGAGCAAGTGGCAATCAGCGTTTATATGCGGAAACTAAG 600
DB 581 TATGCAATCTACCTTCACTCAGCGCTGAAGTAAAGTTTATGAGCAACTCTTG 640
QY 601 AACGTGATACGAGCTAGCATCAGCCTCTTACCAATGCGAGTCTTCAATCTTCATGAT 660

DB 641 AATGAGGCTATCTAGAACCCCTCAAGAGAAATTTGGAGCTTTATATCCAGAT 700
QY 661 GTGCAATTTCTGCCCCGAGAGATACCGTATACCTGTACAGTGTCCATTTCAACCACTCAT 720
DB 701 GTGCACTGTGACCCGAGATGACTTTTACCTTTTCAAGTGTGAGAGATCCCAAGCAT 760
QY 721 ATGAGTGTAGCAGTATGATTAATTTCAATTTCCATTTTCGCGATCTTCGCGGA 780
DB 761 CTGAGTGTGAGAGAGAGACGACCTGGGTACAGTTTACGTTACAGTGTATTTTGGGGGT 820
QY 781 ATCAGTGTCTACCAAAAGATCACTGAAAGAAATCAATGATTTTTCGATATTTTGG 840
DB 821 GTTACGTCCCTTAAAGAGAGAGAGCTTTTCAAGTGAATGATTTCTTAACTACTCTG 880
QY 841 GGTGGGGGAGAGAGAGAGATTTTGGGAGAGAAATCATGATGCTGTGATGAAAGTT 900
DB 881 GGATGGGAGGCGAAGACATGACCTCAGACTGAGGTTGACCTCAAGAAATGAAATTT 940
QY 901 TCAAGATATCCGACACAAATTTGCAAGATTAATTAATGATTTAAGCTGACGAGAGCA 960
DB 941 TCCCGCCCCCTCTGAGAGTGGTAAATTAATGATTTTTCACACTGAGACAAAGCC 1000
QY 961 AATCAATTAATTAATGCGCTACAAATATATGAGGCAAGAGCGGATGACACGT 1020
DB 1001 AATGAGTGAACGCAAGACGATGAAGCTTTACACCAAGTGTCAAGTGTGAGAAACA 1060
QY 1021 GACGCGCTAAGCAATCTGAGTAACTGCTTAATTTGGAATTTGAAGCTCTTCACT 1080
DB 1061 GATGGTTGAGTGTGTTCTTATTAATTAATGATCTGTGAAACAAATCTTTATATATC 1120
QY 1081 CGAGCGTGTGATTT 1097
DB 1121 AACATCAGATGATTT 1137

RESULT 8
US-09-997-333-235
Sequence 235, Application US/09997333
Patent No. 6953836
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC27
CURRENT APPLICATION NUMBER: US/09/997,333
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186	PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065111	PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/067770	PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945	PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910	PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322	PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106	PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827	PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088312	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088734	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858	PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861	PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876	PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105	PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440	PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512	

1	PRIOR FILING DATE: 1998-06-16
2	PRIOR APPLICATION NUMBER: 60/089514
3	PRIOR FILING DATE: 1998-06-16
4	PRIOR APPLICATION NUMBER: 60/089532
5	PRIOR FILING DATE: 1998-06-17
6	PRIOR APPLICATION NUMBER: 60/089538
7	PRIOR FILING DATE: 1998-06-17
8	PRIOR APPLICATION NUMBER: 60/089598
9	PRIOR FILING DATE: 1998-06-17
10	PRIOR APPLICATION NUMBER: 60/089599
11	PRIOR FILING DATE: 1998-06-17
12	PRIOR APPLICATION NUMBER: 60/089600
13	PRIOR FILING DATE: 1998-06-17
14	PRIOR APPLICATION NUMBER: 60/089653
15	PRIOR FILING DATE: 1998-06-17
16	PRIOR APPLICATION NUMBER: 60/089801
17	PRIOR FILING DATE: 1998-06-18
18	PRIOR APPLICATION NUMBER: 60/089807
19	PRIOR FILING DATE: 1998-06-18
20	PRIOR APPLICATION NUMBER: 60/089908
21	PRIOR FILING DATE: 1998-06-18
22	PRIOR APPLICATION NUMBER: 60/089947
23	PRIOR FILING DATE: 1998-06-19
24	PRIOR APPLICATION NUMBER: 60/089948
25	PRIOR FILING DATE: 1998-06-19
26	PRIOR APPLICATION NUMBER: 60/089952
27	PRIOR FILING DATE: 1998-06-19
28	PRIOR APPLICATION NUMBER: 60/090246
29	PRIOR FILING DATE: 1998-06-22
30	PRIOR APPLICATION NUMBER: 60/090252
31	PRIOR FILING DATE: 1998-06-22
32	PRIOR APPLICATION NUMBER: 60/090254
33	PRIOR FILING DATE: 1998-06-22
34	PRIOR APPLICATION NUMBER: 60/090349
35	PRIOR FILING DATE: 1998-06-23
36	PRIOR APPLICATION NUMBER: 60/090355
37	PRIOR FILING DATE: 1998-06-23
38	PRIOR APPLICATION NUMBER: 60/090429
39	PRIOR FILING DATE: 1998-06-24
40	PRIOR APPLICATION NUMBER: 60/090431
41	PRIOR FILING DATE: 1998-06-24
42	PRIOR APPLICATION NUMBER: 60/090435
43	PRIOR FILING DATE: 1998-06-24
44	PRIOR APPLICATION NUMBER: 60/090444
45	PRIOR FILING DATE: 1998-06-24
46	PRIOR APPLICATION NUMBER: 60/090445
47	PRIOR FILING DATE: 1998-06-24
48	PRIOR APPLICATION NUMBER: 60/090472
49	PRIOR FILING DATE: 1998-06-24
50	PRIOR APPLICATION NUMBER: 60/090535
51	PRIOR FILING DATE: 1998-06-24
52	PRIOR APPLICATION NUMBER: 60/090540
53	PRIOR FILING DATE: 1998-06-24
54	PRIOR APPLICATION NUMBER: 60/090542
55	PRIOR FILING DATE: 1998-06-24
56	PRIOR APPLICATION NUMBER: 60/090557
57	PRIOR FILING DATE: 1998-06-24
58	PRIOR APPLICATION NUMBER: 60/090676
59	PRIOR FILING DATE: 1998-06-25
60	PRIOR APPLICATION NUMBER: 60/090678
61	PRIOR FILING DATE: 1998-06-25
62	PRIOR APPLICATION NUMBER: 60/090650
63	PRIOR FILING DATE: 1998-06-25
64	PRIOR APPLICATION NUMBER: 60/090653
65	PRIOR FILING DATE: 1998-06-25
66	PRIOR APPLICATION NUMBER: 60/090656
67	PRIOR FILING DATE: 1998-06-25
68	PRIOR APPLICATION NUMBER: 60/090682
69	PRIOR FILING DATE: 1998-06-26
70	PRIOR APPLICATION NUMBER: 60/090853
71	PRIOR FILING DATE: 1998-06-26

PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1664;
Best Local Similarity 51.8%; Pred. No. 7e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

421 CCTAAGATTGTTGTCAGAGCATGCTGTTCTATTATTGTCCTTATAGATCGTGA
461 CCTCAGGATGTAAAGCTTTACAGAGGGTCGCATCTCTGCCACCGAAGAGAG
481 GGCATTTGAGATATAGTCCACATTTGCACTGCTGCTGCCAAGAACATTTGAC
521 AAACACGTGATGTACCTGCTGGAACATCTGCACTCTCTCCAGAGGAGAGCGAT
541 TATGCAATTTTCACTGTTGAGCAAGTGGCAATTCAGACGTTTAAATGCGGAACTAATG
581 TATGCACTTACGTATCCACAGGCTGAGAGTAAATTTTATGAGCCAACTCTTG
601 AAGTTGATACGATGATCAACGCTCTACCCATGCGAGTCTTCACTTTTCAATGAT
641 AATGAGGCTATCTAGAACCCCTCAAGAAAGAAATTTGGAGCTGTTTATTTCCAGAT
661 GTCGATTTACTGCCCCGAATGACCTTAACCTGTAACCTGTCATTTCAATTTAACAGCTCAT
701 GTGACCTGTGACCCGAGATGACTTTTAACTTTTAAAGTGGAGAGCATCCAAAGCAT
721 ATGAGTGTACGATGATTAATTTCAATTTTAACTTCAATTTTGGCGATCTTGGCGGA
761 CTGGTGTGTCGAGAGACGACACTGGGTACAGTTACGTTACATGATATTTTGGGGGT
781 ATCACTGCTAACAAGATCACTGAAGAAATTCATGATTTTGAATGATTTTGG
821 GTTACTGCTTACGAGAGAGAGGATTTTCAAGTGAATGATTTCTTAACAATCTG
841 GGTGGGGGAGAGAGACGATTTGGGAGAGAAACATGATGCTGATGCTGAAAGT
881 GGATGGGAGGAGAGACGATGACCTCAGACTCAGGTTGACCTCCAAAGATGAAAT
901 TCAAGATATCCGACCAATTTGACGATTAATTTAAATGATTTAGACTCGACGAGGAGAG
941 TCCCGCCCTCTGCTGAATGAGTAAATTTAATGATTTCTTCACTGAGACAAAGGC
961 AATCCAGTTAATTAATGCGCTACAAATTAATGGGCAAGCAAGGCGGATGACAGCT
1001 AATGAGGTGAACGAGAAAGATGAAAGCTTTACACAAAGTGTACGATGAGAGAA
1021 GACGCGCTAAGCAATGAGATTAAGCTGTTAATTTGGAATTTGAGAGCTCTTACACT
1061 GATGGGTAGTAGTGTCTTATTAATTAATGATGTTGGAACAAATCTTTATATATC
1081 CGAGCGCTGCTGATTT 1097
1121 AACATCACAGTGATTT 1137

RESULT 9

US-09-992-598-235
Sequence 235, Application US/09992598
Patent No. 6956108
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Auestin L.
APPLICANT: Kijavln, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355

PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090662
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090663
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 76-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAGAGATTGTTGTCAGAGCATGCTGTTATTAATGTCCTTATAGATCTGTGAA 480
DB 461 CCTCAGGAATGTAAAGCTTTACAGAGGGTGCCTCCTGTTCCCAACGGAACAGAG 520
QY 481 GCACATTTGAGATAATGCTCCACATTTGCACTGTTGTCGCCAAACAATGGAC 540
DB 521 AAACACCTGATGTACCTGCTGGAACAATGTGCATCCCTTCTGCAAGGACGACTGGAT 580
QY 541 TATGCAATTTTCATTGTGAGCAAGTGGCAATCAAGCTTTATGCGGGAATATATG 600
DB 581 TATGCAATTCATGCTACCAAGGCTGAAGTAAAGTTTATATGAGCAACTTTTG 640
QY 601 AACGTGATACGACGTAGATCAGCGCTTACCCATGGAAGTTCATCTTTCATGAT 660

Db 641 AATGGGCTATCTAGAACCCCTCAAGAGAAATATGGGACTGCTTATATTCACGAT 700
Qy 661 GTCCATTTACTCCCGGAAGATACCGTAACTCTGACAGTGTCAATTCACACGATCAT 720
Db 701 GTGACCTGTGATCCGAGAAATGACTTTAACTTTCAAGTGTGAGAGAGATCCCAAGCAT 760
Qy 721 ATGAGTGTAGCGATCGATTAATTTCAATTTCAATTTCCATTTGGCGCATCTTGGCGGGA 780
Db 761 CTGTGTGTGTGGAGAGACGACACTGGGTACAGTTACGTTCACGTGATATTTTGGGGGT 820
Qy 781 ATCAGTGCCTAACAAAGATCACTGAAGAAATCAATGATATTTTGAATGATTTTGG 840
Db 821 GTTACTGCCCTTAAGACAGAGACAGTTTTCAGAGTGAATGATCTCTTAACAACACTAGG 880
Qy 841 GGTGGGGCGAGAGACGACGATTTGGGACGAGAAATCATGATGCTGATGCTGATAAGTT 900
Db 881 GGATGGGGAGGCGAGACGATGACCTCAGACTCAGGGTGTGAGCTCCAAAGAAATGAAAATT 940
Qy 901 TCAAGATATCCGACCAATTTGACGATTAATAATGATTAAGCACTCGACGAGAGGAGC 960
Db 941 TCCCGCCCCCTGCTGAGAGTGGTAAATTAATTAATGATTTTCCACACTAGAGCAAAAGC 1000
Qy 961 AATCCAGTTAATTAATGCGCTACAAATTAATGCGCAACGAGCGCCGATGACAGCT 1020
Db 1001 AATGAGGTGAACGCAAGACGATGAAGCTCTTACACCAAGTGTCAAGATCTGAGAGACA 1060
Qy 1021 GACGCGCTTAAGCAATCTGAAGTATTAAGCTCTGTAATCTGGAATTTGAAGCTCTTACACT 1080
Db 1061 GATGGGTGAAGTATGTTCTTATTAATTAATTAATGATCTGTGGAACACAAATCTTTATATATC 1120
Qy 1081 CGAGCGCTGCTGATTT 1097
Db 1121 AAGATCAGATGATTT 1137

RESULT 10
US-09-949-016-1767
; Sequence 1767, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1767
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1767

Query Match 13.5%; Score 155.4; DB 3; Length 2151;
Best Local Similarity 51.8%; Pred. No. 7.3e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

Qy 421 CCTAAGATTGTGTGCAAGGCAATGCTGTGCTATTAATTTGCCCCCTAATAGAGATGTTAA 480
Db 653 CCTCAGGAATGTAAAGCTTTACAGAGGGTCCGCAATCTGTTCCCAAGCAAGAGAG 712
Qy 481 GACATTTAGAAATATGTTCCCAATTTGCACTGTGCTGCGCAAAACAAATTTGAC 540
Db 713 AAACACCTGATGTACTGCTGGAACATCTGCAATCCCTTCTGCAAGGAGGAGCTGGAT 772

Qy 541 TATGCAATTTTCAATTTGTGAGCAATGCGAATCAGACGTTTAATTCGCGGAAACTAATG 600
Db 773 TATGCAATTTAATGATCCTACAGGCTGAAGTAAATTTAATTCAGCCAAACTCTTG 832
Qy 601 AACGTTGATACGAGTATGATCAAGCTCTTAACCCATGCGAGTGTTCATCTTTCATGAT 660
Db 833 AATGGGCTATCTAGAAAGCCCTCAAGAGAAATTTGGGACTGCTTATTAATTTCCACAT 892
Qy 661 GTCCATTTACTGCCCGGAAGATGACCTTAACCTGTACAGTGTCCAAATTCACACGATCAT 720
Db 893 GTGACCTGTATCCGAGAAATGACTTTAACTTTACAGTGTGAGAGACATCCCAAGCAT 952
Qy 721 ATGAGTGTAGCGATCGATTAATTTCAATTTAATTTCCATTTTGGCGCATCTTGGCGGGA 780
Db 953 CTGTGTGTGTGGAGAGAACAGACAGCTGGTACAGTTACCTTTCAGTGAATATTTTGGGGGT 1012
Qy 781 ATCAGTGCCTAACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTTGG 840
Db 1013 GTTACTGCCCTTAAGACAGACAGATTTTTCAGGTTGAATGATTTCTTAACAACACTAGG 1072
Qy 841 GGTGGGGCGAGAGAGAGATTTTGGCGAGAACATGATGCTGAGCTGAGAAATT 900
Db 1073 GGATGGGAGGCGAGAGACGATGACCTCAGACTCAGGGTTGAGCTCAAAAGATGAAAATT 1132
Qy 901 TCAAGATATCCGACCAATTTGACGATTAATAATGATTAAGCACTCGACGAGAGCGAGC 960
Db 1133 TCCCGCCCCCTGCTGAGAGTGGTAAATTAATTAATGATTTTCCACACTAGAGCAAAAGC 1192
Qy 961 AATCCAGTTAATTAATGCGCTACAAATTAATGCGCAACGAGCGCCGATGACAGCT 1020
Db 1193 AATGAGGTGAACGCAAGACGATGAAGCTCTTACACCAAGTGTCAAGTCTGGAGAAC 1252
Qy 1021 GACGCGCTTAAGCAATCTGAAGTATTAAGCTCTGTAATCTGGAATTTGAAGCTCTTACACT 1080
Db 1253 GATGGGTGAAGTATGTTCTTATTAATTAATTAATGATCTGTGGAACACAAATCTTTATATATC 1312
Qy 1081 CGAGCGCTGCTGATTT 1097
Db 1313 AAGATCAGATGATTT 1329

RESULT 11
US-09-949-016-541
; Sequence 541, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 541
; LENGTH: 2167
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-541

Query Match 13.5%; Score 155.4; DB 3; Length 2167;
Best Local Similarity 51.8%; Pred. No. 7.4e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

Qy 421 CCTAAGATTGTGTGCAAGGCAATGCTGTGCTATTAATTTGCCCCCTAATAGAGATGTTAA 480
Db 653 CCTCAGGAATGTAAAGCTTTACAGAGGGTCCGCAATCTGTTCCCAAGCAAGAGAG 712

RESULT 13
US-09-373-902-4
Sequence 4, Application US/09373902
Patent No. 6649737
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Guegler, Karl J.
Corley, Neil C.
Shah, Puri
Paterson, Chandra
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: 12-Aug-1999
APPLICATION NUMBER: US/09/373,902
CLASSIFICATION: <Unknown>
PRIOR APPLICATION NUMBER: US/09/055,097
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0490 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGTUT06
CLONE: 2551161
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-373-902-4

Query Match 13.5%; Score 155.4; DB 3; Length 2280;
Best Local Similarity 51.8%; Pred. No. 7.6e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTGCAAGAGATGCTGTTGCTATTATGNGCCCTATAGAGATGTAA 480
DB 771 CCTCAGAAATGTAAGACTTTTACAGAGGTCGCCATCTGTTCCCAACCGAAGAGAG 830
QY 481 GCACATTTGAGATATATGCTCCACATTTGCACTGCTGCTGCCAACAACAATTTGAC 540
DB 831 AAACACGTATGTACTGCTGAGACATCTGCACTCTTCTGCAAGAGAGAGAGCTGAT 890
QY 541 TATGCAATTTTCAATGCTGAGCAAGTGGCAATCAGACGTTTAAATGCGGAAACTAATG 600
DB 891 TATGCAATCTACGATCAACACAGGCTGAAGGTAAAGTTTAAATGAGCAAACTCTG 950
QY 601 AACGTTGATAGAGATGATCAAGCCCTTACCAATGAGAGGCTTCACTTTTATGAT 660
DB 951 AATGTTGGCTTACTTAAAGCCCTCAAGAAATAATTTGGAGCTGTTTATTTCCAGAT 1010
QY 661 GTGATTACTGCCGAAAGATGACCGTAACTGTACACGTGTCAATTCACACAGTCAT 720

DB 1011 GTGACCTGTATCCCGAAGATGACTTTTAACTTTAACAAGTGAAGAGACATCCAGCAT 1070
QY 721 ATGAGTATGAGATGATGATTAATTTCAATTAATTAATTTCAATTTGCGGCGCA 780
DB 1071 CTGGTGTGGTGAAGAAAGACATGCGGTACAGTTTACGTTACGTATATTTTGGGGT 1130
QY 781 ATCAGTCACTTAACAAAGATCACCTGAAGAAAAATCAATGATTTTCCAAATGATTTTGG 840
DB 1131 GTTACTGCCCTTAGAGAGAGACAGTTTTCAGAGGAATGATTTCTTAACACTACTG 1190
QY 841 GATTGGGCGGAGAGACGACGATTTGGCGACGAGAACATGATGCTGACCTGAAAGTT 900
DB 1191 GGATGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAGATGAAATTT 1250
QY 901 TCAAGATATCCGACACAAATTTGACGATTAATAATTAAGATCTGACCGAAGCAGC 960
DB 1251 TCCGCGCCCTGCTGAAGTGGTAAATATTAATGATGCTTCCACACTAGAACAAAGGC 1310
QY 961 AATCAGTTTAATTAATGCGGCTACAAATTAATGGGCCAAAGCGCGGATGACAGT 1020
DB 1311 AATGAGTGAACGAAACGATGAGCTTTACACCAAGTGTACAGAGTCTGAGAAC 1370
QY 1021 GACGGCTTAAGCAATCTGAATTAAGCTGTAATCTGGAATTTGAAGCTCTTACACT 1080
DB 1371 GATGGGTTGAGTATGTTCTTATTAATTAATGATCTGTGAACAACAATCCTTATATATC 1430
QY 1081 CGAGCGCTGCTGATTT 1097
DB 1431 AACATCACAGTGATTT 1447

RESULT 14
US-09-949-016-4375
Sequence 4375, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4375
LENGTH: 2965
TYPE: DNA
ORGANISM: Human
US-09-949-016-4375

Query Match 13.0%; Score 150; DB 3; Length 2965;
Best Local Similarity 56.8%; Pred. No. 5.1e-38;
Matches 276; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 405 CGGTGACATGGAATGCGCTTAAGATTGTTGCAAGAGATGCTGTTGCTATTATTTGTC 464
DB 336 CGAGGTACTGGAAGCCTTTGATTTGATGCTCTGGTGAAGGTGGGATCTTATCCC 395
QY 465 CTATAGATGCTGAAGACATTTGAGATATATGCTCCACATTTGCACTGCTGCTGC 524
DB 396 CTTCGGAACGCGACGACACTCCAGTCTGTTGAGACACCTGTTCCATCTCCA 455
QY 525 CAACAAATTTGACTATGCAATTTTATTTGATGTGAGGAAGGGGATTCAGAGCTTAA 584
DB 456 GCGCAGGCGCTGCACTTTTATGTGTGTAACAGTTGGTATCCCAACCTTTAA 515

```
QY 585 TCGGGGAACTATGATGAAGTTGGATACGACGTACGACGCTTACCAATGGCAGTG 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 TCGAGCAAGCTTTTCAACGTTGGCTTTCAAGAAAGCAAGAAAGACTTGGATGGACTG 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 CTTCATCTTTTCATGATGTCGATTACTGCCGGAAGTACCGTAACCTGTACAGCTGCC 704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 TTGATTTTTCATGATGATGATACATACCGGAAGATGATCGCACTATATGATGATGG 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 705 AATTCAACCAAGTCATATGATGATGATGATGATGATGATGATGATGATGATGATG 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 ACAGATGCCGAGGCAATTTTGCACCAAAATGATATGATATGATCTCTTATAC 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 765 GCGCATCTTCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 CGAGTCTTTTGGCGGAGTGTGCTTACAGTGAAGCAATTTGGAAGAAATCAATGGCTT 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 TTGCAATGATTTTGGGGTTGGGGTGGGAGAGACGACGATTTGGCGACGAAATCGAT 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 756 TCTTAATGCTTCTGGGGTTGGGGTGGAGAGATGACGACCTCTGGAACAGATACAGAA 815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 GCGTGG 890
    ||| ||| |||
Db 816 TGCAGG 821
```

```
RESULT 15
US-09-949-016-680
; Sequence 680, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 680
; LENGTH: 4646
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-680
```

```
Query Match 13.0%; Score 150; DB 3; Length 4646;
Best Local Similarity 56.8%; Pred. No. 6.4e-38;
Matches 276; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
```

```
QY 405 CCGTGCATGATGATGCTTAAGATGTTGCAAGGCATCGTGTCTATTATTTGACC 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 CCGAGGTCACTGAAGCTTCTGATTCATGCTCGGTGAAGGTGGCATCTTATCCC 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 465 CTATAGATGCTGAGACACATTTGAGAATPATGCTCCACAATTTGCACTGTTGCTGCC 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 CTTCGGGAACCGCACAGGACCTCCAGTCTGTTCAACACCTGCTTCCCACTCCA 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 525 CAACAACAATTGACTATGCAATTTTCATTTGAGAGAGTGGCAATCAAGAGCTTTAA 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 GCGCCAGCGCTTGCAATTTTATGTTGTTGAACAGTTGTATCCCAACCTTTAA 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 TCGGGGAACTATGATGAAGTTGGATACGATGATGATGATGATGATGATGATGATG 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 TCGAGCATGCTTTTCAACGTTGGCTTTCAAGAGCAATGAAAGACTTGGATGGACTG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 CTTCATCTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 TTGATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 705 AATTCAACCAAGTCATATGATGATGATGATGATGATGATGATGATGATGATGATG 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 ACAGATGCCGAGGCAATTTTGCACCAAAATGATATGATATGATATGATATGATATG 918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 765 GCGCATCTTCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 919 CGAGTCTTTTGGCGGAGTGTGCTTACAGTGAAGCAATTTGGAAGAAATCAATGGCTT 978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 TTGCAATGATTTTGGGGTTGGGGTGGGAGAGACGACGATTTGGGAGAGAAACATCGAT 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 TCTTAATGCTTCTGGGGTTGGGGTGGAGAGATGACGACCTCTGGAACAGATACAGAA 1038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 GCGTGG 890
    ||| ||| |||
Db 1039 TGCAGG 1044
```

```
Search completed: January 8, 2006, 22:30:51
Job time : 247 secs
```


XX Claim 5; Fig 1; 51dp; English.

CC The present sequence encodes a beta1,4-N-acetylglactosaminyltransferase,
 CC designated beta4GalNacT. The enzyme is required for the biosynthesis of
 CC animal cell glycoproteins. The enzyme functions to synthesize the
 CC LacdMac or LDN motif GalNAcbeta4GlcNAc-R. This motif is glycoprotein
 CC hormones produced by the pituitary gland, and is a major marker of
 CC glycoconjugates made by parasitic and non-parasitic invertebrates and may
 CC be implicated in host immune responses to infection. The protein and
 CC polynucleotides are useful for transforming or transfecting host cells
 CC for producing substantially pure forms of the enzyme, or in vitro, for
 CC formation of a LDN structure on proteins or peptides.

CC Sequence 1152 BP; 314 A; 275 C; 259 G; 304 T; 0 U; 0 Other;

Query Match 100.0%; Score 1152; DB 12; Length 1152;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTTTCGTGATTTGGAGTCGCGCAGCTCAAGTCGTTGCTCGTACTTTGTGCGCTT 60
 Db 1 ATGGCTTTTCGTGATTTGGAGTCGCGCAGCTCAAGTCGTTGCTCGTACTTTGTGCGCTT 60
 QY 61 CTTCTATTAGTTCATGCAATGATTTATAGATTCCATCGCTTTACGAGAACCTTACTATC 120
 Db 61 CTTCTATTAGTTCATGCAATGATTTATAGATTCCATCGCTTTACGAGAACCTTACTATC 120
 QY 121 GGCCTCTCGACCTTATTTGCGCAGCTGACGCAATGAGAGCACTGCTCGGGAATTACGGCT 180
 Db 121 GGCCTCTCGACCTTATTTGCGCAGCTGACGCAATGAGAGCACTGCTCGGGAATTACGGCT 180
 QY 181 TCCACTTCGGATGATCTACTTGTAGTCGTGGAATTCACGTTTCCACGATTTCTGAAGTT 240
 Db 181 TCCACTTCGGATGATCTACTTGTAGTCGTGGAATTCACGTTTCCACGATTTCTGAAGTT 240
 QY 241 AATCAGACTAGTTTATGAGAGCAATTCGTCCAACTCTGTTCCCGCAACAACAGACTCTT 300
 Db 241 AATCAGACTAGTTTATGAGAGCAATTCGTCCAACTCTGTTCCCGCAACAACAGACTCTT 300
 QY 301 CAATTCTGTATCAGACACTCCCGCACTGCTGGAGCCATCCGTGATTCCTTCGATGAG 360
 Db 301 CAATTCTGTATCAGACACTCCCGCACTGCTGGAGCCATCCGTGATTCCTTCGATGAG 360
 QY 361 CCGGACTTCAAACTCTCGAAGAAATCTATCCGACACGACGCGGTGACATGGAATG 420
 Db 361 CCGGACTTCAAACTCTCGAAGAAATCTATCCGACACGACGCGGTGACATGGAATG 420
 QY 421 CCTAAGGATTGTGTGCAAGGCAATCGTGTGCTATTATTTGTCCTATAGATCGTGAA 480
 Db 421 CCTAAGGATTGTGTGCAAGGCAATCGTGTGCTATTATTTGTCCTATAGATCGTGAA 480
 QY 481 GCACATTGGAATAATGCTCCACAAATTTGCACTCGTTCGCGCAAAACAATTTGAC 540
 Db 481 GCACATTGGAATAATGCTCCACAAATTTGCACTCGTTCGCGCAAAACAATTTGAC 540
 QY 541 TATGCAATTTTCATTTGGAAGAGGAGGCGAATGAGAGCTTTAATGCGGGAAATCAATG 600
 Db 541 TATGCAATTTTCATTTGGAAGAGGAGGCGAATGAGAGCTTTAATGCGGGAAATCAATG 600
 QY 601 AACGTTGATACGACGTAGCATCGCCTCTACCCATGAGAGTGCTTCATCTTTCATGAT 660
 Db 601 AACGTTGATACGACGTAGCATCGCCTCTACCCATGAGAGTGCTTCATCTTTCATGAT 660
 QY 661 GTGCAATTTACTGCCGGAAGATGACCGTAACTGTACACGTGTCCAAATTCACACAGTCAT 720
 Db 661 GTGCAATTTACTGCCGGAAGATGACCGTAACTGTACACGTGTCCAAATTCACACAGTCAT 720
 QY 721 ATGAGTGAAGGATGCAATTAATCAATTAATTCATTCATATTTGGGAGATCTTCGGGGA 780
 Db 721 ATGAGTGAAGGATGCAATTAATCAATTAATTCATTCATATTTGGGAGATCTTCGGGGA 780
 QY 781 ATCAGTGCACCTAACAAAAGATCACTGAAAGAAAATCAATGATTTTCGAATGATTTTGG 840

Db 781 ATCAGTGCACCTAACAAAAGATCACTGAAAGAAAATCAATGATTTTCGAATGATTTTGG 840
 QY 841 GGTGGGGGGGAGAGAGAGAGATTTGGCGAGACAGAACATGATGGCTGAGCTGAAAGTT 900
 Db 841 GGTGGGGGGGAGAGAGAGAGATTTGGCGAGAGACATGATGGCTGAGCTGAAAGTT 900
 QY 901 TCAAGATATCCGACACAAATTGCGACGATATTAATGATTTAAGCACTCGACGGAGCGACG 960
 Db 901 TCAAGATATCCGACACAAATTGCGACGATATTAATGATTTAAGCACTCGACGGAGCGACG 960
 QY 961 AATCCAGTTAATTAATGCCGCTCAAAATATATGGCCCAAGCAGACGCGGATGACAGCT 1020
 Db 961 AATCCAGTTAATTAATGCCGCTCAAAATATATGGCCCAAGCAGACGCGGATGACAGCT 1020
 QY 1021 GACGGCTTAAGCAATCTGAAGTAAAGCTCGTAAATCTGGAATTTGAAGCTCTTACACT 1080
 Db 1021 GACGGCTTAAGCAATCTGAAGTAAAGCTCGTAAATCTGGAATTTGAAGCTCTTACACT 1080
 QY 1081 CGAGCCGTCGTGATTTGCTCGAAAAAGACTCCGCGGAGACTCGAAGGAGACTTTTCCA 1140
 Db 1081 CGAGCCGTCGTGATTTGCTCGAAAAAGACTCCGCGGAGACTCGAAGGAGACTTTTCCA 1140
 QY 1141 ACGTGTTTTAA 1152
 Db 1141 ACGTGTTTTAA 1152

RESULT 2
 ABL14209
 ID ABL14209 standard; cDNA, 1839 BP.
 XX
 AC ABL14209;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37109.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR P-P-SDB; ABB70106.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 37109; 21dp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABJ30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
 CC from WFO at ftp.wfo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1839 BP; 501 A; 450 C; 384 G; 504 T; 0 U; 0 Other;

Query Match 16.7%; Score 192.6; DB 4; Length 1839;
 Best Local Similarity 58.7%; Pred. No. 9.9e-52;
 Matches 333; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

399 GCACCCCGGTGACATGGAATGCTTAAGATTGTGTGAGAGGATCGGTGTCTATTAT 458
 DB 600 GCGCCCTGTGGCGGCTTCGACGCTGAAAACGCAATGCCGACATCACGTGGCTATTGT 659
 OY 459 TGTGCTCATATAGATTCGTAAGACATTTGGAATTAATGCTCCCAATTTGACTCGTT 518
 DB 660 TGTGCTCTCCGCAATCGATACGCGCCATTTATTCTTTCTGCGCAACATCACCCATT 719
 OY 519 GCTCCCAACACCAATTTGACTGCAATTTTCAATTTGAGAGCAAGTGGCGAATCAGAC 578
 DB 720 TCTGATGAAGAGCCGATCGCTTATCGCATTTTCAATTTGAGAGCAAGCAACGGGAAGCC 779
 OY 579 GTTTAATGCGGGAACATTAATGAGTTGATACGACGTAGATCACGCTTACCCATG 638
 DB 780 CTTTATCGGCGTCCATGATGAATGATGTTTGAAGGCTTAAAGCTGTACCGAGT 839
 OY 639 GCAGTCTTCATCTTTCATGATGATGATTTACTGCGCCGAAGTGAACCGTAACCTGTAAC 698
 DB 840 GGAATGTTTATATTCACGATGTGATTTTCTGCTTTGAGACGACCGCAATCTTTACAA 899
 OY 699 GTGTCATATTAACACGCTCATATGATGATGATGATGATGATGATGATGATGATGAT 758
 DB 900 CTGTCCACGTGACGCGCGACACATGTCAGTGTGCTATAGACAGTGAACCTTACAGTTGCC 959
 OY 759 ATATTGCGGATCTTTCGCGGGAATCAGTGCATTAACAAAGTCACTCGAAGAAATCAA 818
 DB 960 TTATGATCAATATTTGAGAGGTGTTTCGCAATGACGCGTGAAGCACTTTCAGCCGTAA 1019
 OY 819 TGGATTTTGAATGATTTTGGGGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
 DB 1020 TGGGTTCTCAAACTGCTTCTTGGCTGGGGCGGCGAGATGACGACATGTCCAAACGGTT 1079
 OY 879 ATCGATGCTGAGCTGAAGAGTTTCAAGATATCGACACAAATGTCAGATATTAATGAT 938
 DB 1080 GAAGACGGAACCTATTCATATCAAGGTATCCGGTCAATAGCTCGCTACAAGATGCT 1139
 OY 939 TAAGCATCTGACGAGAGCGACGATCC 965
 DB 1140 GAAGCATCGAAGAGAAAGGCCAATCC 1166

RESULT 3
 AA242120
 ID AA242120 standard; cDNA; 1938 BP.

AC AA242120;
 DT 31-JAN-2000 (first entry)
 DE Human endometrium tumour cDNA derived EST 140.
 XX
 XX Endometrium; human; tumour; cancer; anticancer; cytostatic;
 XX EST: treatment; uterine; gene therapy; expressed sequence tag; ss.
 OS Homo sapiens.

XX
 XX DE19817948-A1.
 XX
 XX 21-OCT-1999.
 XX
 XX 17-APR-1998; 98DE-01017948.
 XX
 XX 17-APR-1998; 98DE-01017948.
 XX

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX WPI; 1999-591957/51.
 DR
 XX
 XX
 PT New nucleic acid sequences expressed in uterine cancer tissues, and
 PT derived polypeptides, for treatment of uterine and endometrial cancer and
 PT identification of therapeutic agents.

Claim 3; Page 273; 444pp; German.

This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 CC are used (i) to identify agents suitable for treatment of uterine or
 CC endometrial cancer; (ii) directly for treating these forms of cancer
 CC (including expression from gene therapy vectors) and (iii) for generation
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent different
 CC parts of the same unknown gene, distorting the estimated frequency of
 CC occurrence in a particular tissue. AA241981-242121 represent EST
 CC fragments derived from a human endometrium tumour cDNA library which
 CC encode the protein sequences represented in AA59941-Y60328

Sequence 1938 BP; 395 A; 567 C; 515 G; 461 T; 0 U; 0 Other;

Query Match 14.5%; Score 167.6; DB 2; Length 1938;
 Best Local Similarity 54.4%; Pred. No. 1.7e-43;
 Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CTTAAGATGTTGTTCAAGGACATGTTGCTATTATTGCTTATAGAGATGTGA 480
 DB 573 COTGAGGTTGTGAGCCCCGCTCCGGAACAGCATATGTCCTCATGTCGCCGGAG 632
 OY 481 GCACATTTGAATATATCTCAACAATTTGACATGTCGTCGCCAACAACAATTTGAC 540
 DB 633 CACCACTCGCCCTGCTCTTACCACTGCAACCCCTTTCAGACGCGCAGAGCTTGT 692
 OY 541 TATGCAATTTTCATGTTGAGAGAGTGGAGATGACAGCTTAAATCGCGGAACTAATG 600
 DB 693 TATGCAATCTATGATATCACCGAGCTGAATGGAACATTTTACAGGCAAAATCTTGG 752
 OY 601 AACGTTGATACGACGTAGATCAAGCTCTTACCCATGAGCGAGTCTTATGATGAT 660
 DB 753 AACGTTGGGGTGGAGAGGCGCTGCTGATGAAGATGGAGCTGCTGTTCTGACAGAT 812
 OY 661 GTGATTTACTGCCCCGAAGATGAGCCGTAACTGTACAGTGT---CCAATTACACAGT 717
 DB 813 GTGACCTCTTGGCCAGAAATGACCAACATCTGTATGTGTGTGACCCCGGGAGCCGCG 872
 OY 718 CATATGAGTGTAGAGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 777
 DB 873 CATGTTGCCGTTGCTATGAACAAGTTTGGATACAGCTCCGTAACCCCAAGTCTTGGGA 932
 OY 778 GGAATCAGTGCATTAACAAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTT 837
 DB 933 GGAATCTAGACCTTAACCTGACCAAGTACCTGAAGATGATGATGATGATGATGATGAT 992
 OY 838 TGGGTTGGGGCGAGAGAGAGAGATTTGGCGAGAGAAACATCGATGGCTGGAGCTGAAA 897
 DB 993 TGGGGCTGGGGTGGAGAGATGACACATTTGCTACAGAGTGGCGCTGGCTGGAGTAAAG 1052
 OY 898 GTTCAAGATATCCGACCAATTTGACGATATTAATTAATTAATTAATTAATTAATTAATTA 957
 DB 1053 ATCTCTGAGCCCCCAACATCTGTAGAGACATTAATTAATTAATTAATTAATTAATTA 1112
 OY 958 ACGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1017

DB 1113 GGCAATGAGAAAATCCACAGATTGACCTTCGTCGCTACCCAGAAATTCCTGAGC 1172
QY 1018 CGTGACGGCCTTAAGCATCTGAAGTATTAAGTCTGAATTTGAAGCCTCTCTAC 1077
DB 1173 CAAGATGGAGATGAATCTACCTGACATACCAATGCTGCTGAGAGCTGGGGCTTTTAT 1232
QY 1078 AC 1079
DB 1233 AC 1234
RESULT 4
ACN37523
ID ACN37523 standard; cDNA; 1949 BP.
XX ACN37523;
XX 18-NOV-2004 (first entry)
XX Tumour-associated antigenic target (TAT) cDNA DNA323958, SEQ ID NO:468.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX tumour; diagnosis; cell proliferative disorder; breast cancer;
XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX central nervous system cancer; bladder cancer; pancreatic cancer;
XX cervical cancer; melanoma; leukemia; hybridisation probe;
XX chromosome identification; chromosome mapping; gene mapping;
XX gene therapy; cytostatic; gene; ss.
XX Homo sapiens.
XX WO2004030615-A2.
XX 15-APR-2004.
XX 29-SEP-2003; 2003WO-US028547.
XX 02-OCT-2002; 2002US-0414971P.
XX (GETH) GENENTECH INC.
XX Wu TD, Zhang Z, Zhou Y,
XX
XX MPI; 2004-347921/32.
XX P-PsDB; ABM80189.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
XX useful in preparing a medicament for treating or detecting a
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX prostate cancer or tumor.
XX
XX Claim 1; SEQ ID NO 468; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acids and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention

XX SQ Sequence 1949 BP; 376 A; 584 C; 525 G; 464 T; 0 U; 0 Other;
Query Match 14.5%; Score 167.6; DB 13; Length 1949;
Best Local Similarity 54.4%; Pred. No. 1.7e-43;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;
QY 421 CCTAAGATTTGTGTGCAAGGCGATCGTGTGCTATTTATTTGTCCTTATAGATCGTGA 480
DB 608 CCTGCAGGTTGTGAAGCCCGGCTCCGAAACAGCATATTTGCTCATGTGCTCCGGAG 667
QY 481 GCACATTTGAGATATATGCTCCACAAATTTGCACTGTTGCTGCCAACAACATTTGAC 540
DB 668 CACCACTGCGGCTGCTGCTTACCACTGCACCCCTTTCACAGCCGAGAGCTTGT 727
QY 541 TATGCAATTTTCATTTGAGAGCAATGGGCAATCAAGCTTTAATCGCGGAAACTAAT 600
DB 728 TATGGCATTTATGTCATTCACCAAGCTGGAATTTGAACATTTAACAGGCAAACTGTTG 787
QY 601 AACGTTGATPACAGATACATCACGCTCTTACCCATGCGAGTGTTCATCTTTCATGAT 660
DB 788 AACGTTGGGGTGGAGAGGCCCTGGGTGATGAAGGTGGACTGGCTGTTTTCAGCAT 847
QY 661 GTGATTTACTGCCGAGAGATGACGTTAAGCTTACAGCTGT---CCAATTCAACGAGT 717
DB 848 GTGACCTCTTGCAGAAATGACCAATCTGTATGTGTGAGACCCCGGGGACCCCGC 907
QY 718 CATATGAGTGTACGATGATTAATTCATTTAATTCATATTGCGGATCTTCCGCG 777
DB 908 CAGTTGCCGTTGCTATGAAACAAAGTTTGATACAGCTTCCGTAACCCAGTACTTCGA 967
QY 778 GGAATCAGTGCATCAAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTT 837
DB 968 GGAATCTCAGCACTTACTCTGACGATCCTGAAGATGAATGGCTTCCCATGATGATAC 1027
QY 838 TGGGCTTGGGGCGAGAGGACGACGATTTGGGACGAGAACATCATGCTGTGACTGAAA 897
DB 1028 TGGGCTGGGGTGTGAGGATGACGACATTTGCTACCGAGGTGCGCTGTGGGATGAA 1087
QY 898 GTTTCAGATATTCGACACAAATTTGACGATTAATTAATTAAGCACTGCAGGAAGCG 957
DB 1088 ATCTCTGGCCCCCAGACTCTGTAGGACACTTAAATGTGAAGACCGAGAGATTAAG 1147
QY 958 ACGAATCCAGTTAATTAATGCGCTTCAAAATTAATGGGCCAAACGAAGCCGATGAGAC 1017
DB 1148 GGCAATGAGAAAATCCACAGATTGACCTTCGTCGCTGACCAAGATTTCTGAGACG 1207
QY 1018 CGTGACGGCCTTAAGCATCTGAAGTATTAAGTCTGTAATCTGGAATTTGAAGCTCTCTAC 1077
DB 1208 CAAGATGGAGATGAATCACTGACATACAGTTGCTGGCTGAGAGCTGGGGCTTTTAT 1267
QY 1078 AC 1079
DB 1268 AC 1269
RESULT 5
AAC77836
ID AAC77836 standard; cDNA; 2002 BP.
XX AAC77836;
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated gene sequence SEQ ID NO:230.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
XX antidiabetic; antihypertensive; antineoplastic; antiarthritic; antiviral;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
XX vasotropic; antihypertensive; angiogenesis; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;

PT New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.

XX Claim 20; SEQ ID NO 287; 800pp; English.

CC The invention relates to novel isolated polypeptides, mature forms of
 CC these, or a sequence that is at least 95 % identical to, or having one or
 CC more conservative amino acid substitutions in the polypeptides. The
 CC polypeptides, nucleic acid molecules and antibodies are useful in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease, preferably a NOVX-associated disorder. The nucleic acid
 CC molecules, polypeptides and antibodies are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
 CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
 CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
 CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
 CC dyslipidemias. The nucleic acids and polypeptides may also be used as
 CC targets for the identification of small molecules that modulate or
 CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
 CC hemolipolysis, wound healing and angiogenesis, in gene therapy, in
 CC generation of antibodies that bind immunospecifically to NOVX substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridization probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. This sequence
 CC corresponds to the gene encoding one of the NOVX polypeptides of the
 CC invention.

XX Sequence 1062 BP; 303 A; 224 C; 231 G; 304 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 10; Length 1062;
 Best Local Similarity 54.3%; Pred. No. 5.1e-41;
 Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGACCGCGTGACATGATGCTTGAAGATTGTGCAAGCATCTG 448
 DB 297 ATTTAGATATTGAGCGAGGGGTCATTGAGCGCAAAAGCTGTAAACCAATGGAAG 356
 QY 449 TTGCTATTATTGTCCTTATAGATCGTGAACACATTGGAATATCTCCACATT 508
 DB 357 TGGCAGTTCTCATTCCTTCCGTAATCGCATGAAACATCTCCAAATTTTCTTACATC 416
 QY 509 TGCACCTGTTGCTCCGCAAAACAATTTGACATATGCAATTTTCTGAGCAAGTGG 568
 DB 417 TCAATTCAGATGCTCCAGAAAGACGCGCTGAAATTCGTTTATATGTCATTGAACAGACTG 476
 QY 569 CGAATCAGAGCTTTTATCCGCGGAAACTATAGAACCTTGGATACGAGTACATCAGCC 628
 DB 477 GCAACACACCTTTTAAACCTGCGATGCTTTTCAATGTGGCTTCAAAAGCCATGAAG 536
 QY 629 TCTACCATGGACAGTCTTATCTTTCATGATGTCGATTACTGCCGAAGATGACCGTA 688
 DB 537 ACAGTGTCTGGACAGTGTATATCTTCCAGATGTGATCATCTCAAGAAATGACCGGA 596
 QY 689 ACCTGACAGCTGTCCAAATTCACACGTCATATAGTGTAGAGATCGCTAAATTTCAAT 748
 DB 597 ACTATTTACGAGTGTGAGAAATGCAACGTCATTTTGTCTCAAAAGCTGATTAATACATGT 656
 QY 749 ATTAATCTTCATATTTGCGGATCTTGGCGGAATCAGTCACTAACAAAGATTCACCTGA 808
 DB 657 AATATCTTCATATAAGAAATTTTGTGTGTAGTGTGGCTGACAGTGTGAACAATTTA 716
 QY 809 AGAAAAATCATATGATTTTTCGATGATTTTGGGGTGGGCGGAGAGACGAGATTTGG 868
 DB 717 GAAAGATCATATGATTTTCTTAATGCTTCTGGGAGTGGGAGGAGATGATGACCTTT 776
 QY 869 CGACGAGACATCGATGAGTGAAGTTTCAAGATATCCGACACAATTTGACGAT 928
 DB 777 GGAACAGAGTTCACTATGCTGTGATATATGTAACGACCAAGAGGAGACTTAAGAAAT 836

QY 929 ATTAATGATTTAGACTCGACGGAAGCGAATTCAGTTATTAATCCGCTACAAA 988
 DB 837 ACAAGTCAATTTCTCTA-TTACATAGAGTGAAGTCAAGTT--TTAGACCGTATTAAT 893
 QY 989 TTAATGGCCAAACGAAGCCCGCATGACACGTGACGCTTAAGCAATCTGAAGTATAC 1048
 DB 894 TACTAAGGATTTCCAAAGGAGCGTCAGTACATGATGAGTGAACAATTTAATATAGGC 953
 QY 1049 TCGTAATCTGG 1060
 DB 954 CAAAATTAAGTGG 965

RESULT 8

AAV69749 standard; cDNA; 1146 BP.

AAV69749;

04-FEB-1999 (first entry)

Human lactosyl ceramide synthase encoding cDNA.

lactosyl ceramide synthase; enzyme; recombinant; rat; mouse; human;
 uridine phosphate-galactose; glycosyl ceramide; sphingoglycolipid; ds.

Homo sapiens.

Key

CDS

Location/Qualifiers

1..1146

/*tag= a

/product= "Lactosyl ceramide synthase"

/note= "the stop codon is not indicated"

JP10295371-A.

10-NOV-1998.

25-FEB-1998; 96UP-00043335.

28-FEB-1997; 97JP-00045401.

(KAOS) KAO CORP.

(RIKA) RIKAKAKU KENKYUSHO.

WPI; 1999-038272/04.

P-PSDB; AAW81569.

Lactosyl ceramide synthase and its gene - used for recombinant production

of the enzyme.

Claim 4; Page 13; 15pp; Japanese.

This cDNA encodes a human lactosyl ceramide synthase. The invention

provides rat, mouse and human lactosyl ceramide synthases (AAW81567 to

AAW81569) and genes (AAV69749) encoding the enzymes

respectively. The nucleic acids, and host cells transformed with a

recombinant DNA containing the lactosyl ceramide synthase gene are used

for the recombinant production of the enzyme. The new lactosyl ceramide

synthase has the following enzymatic properties: it transfers galactose

from uridine phosphate-galactose to glycosyl ceramide to synthesize

lactosyl ceramide; it reacts strongly with glycosyl ceramide and weakly

with globoside and does not substantially with ceramide, lactosyl

ceramide, GA2 and GM2. Lactosyl ceramide is used as a raw material for

sphingoglycolipid

Sequence 1146 BP; 339 A; 240 C; 242 G; 325 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 2; Length 1146;
 Best Local Similarity 54.3%; Pred. No. 5.3e-41;
 Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

389 ATCCGACACGACCGCGTGACATGATGCTTGAAGATTGTGCAAGCATCTG 448

Db 413 ATTTAGATATTGAGCAGGGGGGTCATTGGAGGCCAAGAGCTGTAAACCAAGATGGAAG 472
 Qy 449 TTGCTATTATTGTCGCCCTATTAGATCGTAGACCAATTGGAGATTAAATGCTCCACAATT 508
 Db 473 TGGCAGTTCTCATCTTCTCCGTAATGCGCATGAACATCTTCCAAATTTTCTTCAATC 532
 Qy 509 TGCACCTGTTGCTCGGCCAAACAAACAATTGACTATGACAAATTTTCAATTGAGACAGTGG 568
 Db 533 TGAATTCAAATGCTCCAGAAAGACGGCTGGAATTTGCGTTTATGTCATTGAACACACTG 592
 Qy 569 CGAATCAGACGTTTAATCGCGGGAACCTAATGAACGTTGATACGAGATCAGACGCC 628
 Db 593 GGACACAACCTTTTAAACGTCGATGCTTTTCATGTGGCTTCAAGAGGCCATGAAG 652
 Qy 629 TCTACCCATGAGCAGTCTCATCTTTTCATGATGTCGATTTACTGCCCCGAAGATGACCGTA 688
 Db 653 ACAGTGTCTGGGACTGTGTATCTTCCACGATGTGATCATCTACCTGAAGATGACCGGA 712
 Qy 689 ACCGTACACGTTGTCGAATTCACACAGTCATATGATGATGAGGATCGATCAATTCATTT 748
 Db 713 ACTATTACGATGTGAGAAATGCCAGCTATTTTGTCTGCAAGCTGATTAATACATGT 772
 Qy 749 ATTAACCTTCATATTGCGCGATCTTGGCGGAATCAGTGCACTAACAAAGATCACCTGA 808
 Db 773 ATATTCTTCATATAAGAAATTTTGTGTGTAAAGTGGCTGACAGTGAACAAATTTA 832
 Qy 809 AGAAATTCATGATGATTTTGCATGATTTTGGGGTGGGGCGGAGAGACGACGATTTGG 868
 Db 833 GAAAGATCAATGTTTCTTAATGCCCTTCTGGGGATGGGGAGAGAAATGATGACCTTT 892
 Qy 869 CGACGAAACATGATGATGCTGAGCTGAGCTGAAAGTTTCAAGATATCGACACAAATTGACGAT 928
 Db 893 GGAACGAGTTCCTATGCTGATATATATTAACACAGACAGAGGAGACTTGAAGAAAT 952
 Qy 929 ATTAATGATTAAGCACTCGACGGAACGACGAATCCAGTTAAATGAATGCGCTACAAA 988
 Db 953 ACAAGTCAATCTCTCA-TCAACATAGAGTGAAGTCCAGTT--TTTGAAGCGGTATTAAT 1009
 Qy 989 TAATGGGCCAAAGAACGCCGATGAGACAGTACGCGCTTAAGCAATCTGAATTAAGC 1048
 Db 1010 TACTTAAGGTATTCACAGAGCGTCAGATCATGATGAGTGAACAAATTAATATATAGGC 1069
 Qy 1049 TCGTAATCTGG 1060
 Db 1070 CAAAATACTGG 1081
 RESULT 9
 ADJ95057 standard; DNA; 1201 BP.
 AC ADJ95057;
 DT 06-MAY-2004 (first entry)
 XX Novel NOVX gene sequence #143.
 XX de; gene; antidiabetic; anorectic; cardiast; hypotensive;
 XX antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
 XX protocoacide; nootropic; neuroprotective; antiparkinsonian;
 XX anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
 XX dermatological; antiaesthetic; antilipemic; gene therapy;
 XX metabolic disorder; diabetes; obesity; infectious diseases; anorexia;
 XX cancer; cardiovascular diseases; hypertension; atherosclerosis;
 XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 XX epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;
 XX inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;
 XX cell differentiation; cell proliferation; hematopoiesis; wound healing;
 XX angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
 OS Homo sapiens.
 XX , XX

PN WO2003040325-A2.
 XX 15-MAY-2003.
 PD 05-NOV-2002; 2002MO-US035464.
 XX 05-NOV-2001; 2001US-033626P.
 XX 06-NOV-2001; 2001US-033072P.
 PR 09-NOV-2001; 2001US-0348283P.
 PR 15-NOV-2001; 2001US-0335610P.
 PR 16-NOV-2001; 2001US-0338543P.
 PR 20-NOV-2001; 2001US-0331630P.
 PR 20-NOV-2001; 2001US-0331641P.
 PR 21-NOV-2001; 2001US-0332152P.
 PR 21-NOV-2001; 2001US-0333461P.
 PR 28-NOV-2001; 2001US-0333912P.
 PR 28-NOV-2001; 2001US-0334027P.
 PR 29-NOV-2001; 2001US-0334300P.
 PR 30-NOV-2001; 2001US-0334421P.
 PR 30-NOV-2001; 2001US-0334526P.
 PR 04-DEC-2001; 2001US-0336576P.
 PR 04-DEC-2001; 2001US-033664P.
 PR 07-DEC-2001; 2001US-0338314P.
 PR 07-DEC-2001; 2001US-0338390P.
 PR 10-DEC-2001; 2001US-0339006P.
 PR 10-DEC-2001; 2001US-0339008P.
 PR 11-DEC-2001; 2001US-0339286P.
 PR 01-FEB-2002; 2002US-0353280P.
 PR 01-FEB-2002; 2002US-0353288P.
 PR 01-FEB-2002; 2002US-0354392P.
 PR 04-FEB-2002; 2002US-0354393P.
 PR 04-FEB-2002; 2002US-0354409P.
 PR 27-FEB-2002; 2002US-0359944P.
 PR 27-FEB-2002; 2002US-0360148P.
 PR 05-MAR-2002; 2002US-0361790P.
 PR 05-MAR-2002; 2002US-0361833P.
 PR 05-MAR-2002; 2002US-0361925P.
 PR 05-MAR-2002; 2002US-0362230P.
 PR 05-MAR-2002; 2002US-0362625P.
 PR 13-MAR-2002; 2002US-0364000P.
 PR 13-MAR-2002; 2002US-0364181P.
 PR 13-MAR-2002; 2002US-0364182P.
 PR 13-MAR-2002; 2002US-0364197P.
 PR 13-MAR-2002; 2002US-0364227P.
 PR 17-MAY-2002; 2002US-0381621P.
 PR 28-MAY-2002; 2002US-0383675P.
 PR 17-JUL-2002; 2002US-0396703P.
 PR 06-AUG-2002; 2002US-0401552P.
 PR 07-AUG-2002; 2002US-0401594P.
 PR 07-AUG-2002; 2002US-0401787P.
 PR 15-AUG-2002; 2002US-0403619P.
 PR 20-AUG-2002; 2002US-0404821P.
 PR 23-AUG-2002; 2002US-0405368P.
 PR 23-AUG-2002; 2002US-0405402P.
 PR 23-AUG-2002; 2002US-0405496P.
 PR 23-AUG-2002; 2002US-0405631P.
 PR 26-AUG-2002; 2002US-0406125P.
 PR 04-NOV-2002; 2002US-00287226.
 XX (CURA-) CURAGEN CORP.
 XX Agee ML, Alsbrook JP, Berghe C, Boldog FL, Burgess CE, Chant JS;
 XX Chaudhuri A, Diplopo VA, Edinger SR, Bisen A, Ellerman K;
 XX Gangoli EA, Gorman L, Gerlach VR, Ji W, Kekuda R, Khramsov NV;
 XX Li L, Malyanar UM, MacDougall JR, Mezes PS, Miller CE, Millet I;
 XX Ooi CE, Ort T, Padigaru M, Patlurajan M, Rastelli L, Rieger DK;
 XX Rothenberg ME, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ;
 XX Verneer CM, Zerhusen BD, Zhong M;
 DR WPI; 2003-441551/41.
 DR P-PSDB; ADJ95058.
 XX New isolated NOVX polypeptides and polynucleotides, useful for

preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

Claim 20; SEQ ID NO 285; 800bp; English.

The invention relates to novel isolated polypeptides, mature forms of these, or a sequence that is at least 95 % identical to, or having one or more conservative amino acid substitutions in the polypeptides. The polypeptides, nucleic acid molecules and antibodies are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder. The nucleic acid molecules, polypeptides and antibodies are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases (hypertension, atherosclerosis), neurodegenerative disorders (Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders (osteoarthritis)), hematoepietic disorders, inflammatory skin disorders, asthma, and various dyslipidemias. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, hematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOVX substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridization probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. This sequence corresponds to the gene encoding one of the NOVX polypeptides of the invention.

Sequence 1201 BP; 355 A; 250 C; 262 G; 334 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 10; Length 1201;

Best Local Similarity 54.3%; Pred. No. 5.4e-41; Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

389 ATCCGACAGCAGCCGCGTGACATGATGCTTAAGATTGTGTGCAAGCATCTGTG 448
436 ATTTAGATATTGAGCAGGGGTCATTGGAGCCAAAGACCTTAACCAATGGAAG 495
449 TTGCTATTATTGTGCTTATAGATCGTGAACACATTTGAGATATGCTCCCAATT 508
496 TGGCAGTTCTCATTTCTTCCGTAATCGCATGAACATCTCCCAATTTTCTTCAATC 555
509 TGCACGCTGTGTCGCGCAACAACTGATGATGATTTTCAATTGAGCAAGTGG 568
556 TTAATTCAGATGCTCCAGAAAGCAGCGCTGAATTTGCTTTATGTCATTGAACAGACTG 615
569 CGAATCAGACGTTTATATCGCGGAAACTATGAACTGATGATGATGATGATGATGATG 628
616 GCACACAACTTTTAAACGTGCGATGCTTTTCAATGTGGCTTCAAGAGCCCATGAAG 675
629 TTTACCCATGCAAGCTTCTTATCTTTCATGATGATGATGATGATGATGATGATGATG 688
676 ACAGTGTCTGGGACGTGTATATCTTCCAGATGATGATGATGATGATGATGATGATG 735
689 AACTGACACGCTGTCATTTCAACCACTGATGATGATGATGATGATGATGATGATGATG 748
736 ACTATTTAGGATGAGAAATGCAAGCTGATTTGCTTCAAAAGCTGAGTAAATATCATGT 795
749 ATAACTTCATATTCGCGATCTTTCGCGGATCATGCTACATCAAAAGATTCACCTGA 808
796 ATATTTCTTCATATAAGAAATTTTGGTGGTGAAGTGGGCTGACAGTGAACAATTTA 855
809 AGAAATCATGATTTTGAATGATTTTGGGCTGGGCGAGAGAGACGATTTGG 868
856 GAAAGTCATGATGTTTCTTAATGCTTCTGCGGATGCGGAGAGAGATGATGATGATG 915
869 CGACGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
916 GGAACGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 975
939 ATAAATGATTAAGCATCGACGAGAGCAATTCAGTTAATATGCTCCGCTACAAAA 988

Db 976 ACAAGTCATTTCTCTCA-TCACCATGAGGTGAATCCAGTT--TTAGGACGGTATTAAT 1032
Qy 989 TATAGGCGCAACGAAAGCCCGATGACACGTCGAGCGCTTAAGCATCTGAAGTATTAAC 1048
Db 1033 TACTAAGGTATTCCAAAGAGCGTCAGTACATGATGATGATGATGATGATGATGATGATG 1092
Qy 1049 TCGTAATCTGG 1060
Db 1093 CAAAATTAATCTGG 1104

RESULT 10

ADQ89159
ID ADQ89159 standard; cDNA; 3931 BP.

AC ADQ89159;

DT 21-OCT-2004 (first entry)

DE Human urological disorder related protein 21165 encoding cDNA SEQ:111.

KW urological disorder; uropathic; cytostatic; urinary incontinence;

XX benign prostatic hyperplasia; human; gene; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS /tag= a
/product= "urological disorder related protein 21165"

PN MO2004065576-A2.

XX 05-AUG-2004.

XX 14-JAN-2004; 2004WO-US000750.

XX 15-JAN-2003; 2003US-0440318P.

XX 04-FEB-2003; 2003US-0444783P.

XX 27-MAR-2003; 2003US-0457901P.

XX 08-MAY-2003; 2003US-0468775P.

XX 19-MAY-2003; 2003US-0471614P.

XX 16-JUN-2003; 2003US-0478742P.

XX 18-JUL-2003; 2003US-0488529P.

XX 30-UTL-2003; 2003US-0491156P.

XX 02-SEP-2003; 2003US-0499594P.

XX 26-SEP-2003; 2003US-0506332P.

XX (MILL-) MILENNIUM PHARM INC.

XX Karicheti V, Silos-Santiago I, Eliasof SD;

XX WPI; 2004-562167/54.

XX P-PSDB; ADQ89160.

XX Use of polypeptides related to urological disorders, e.g. 44190, 54181,

XX 211 or for identifying a compound capable of treating a urological

XX disorder or identifying and treating a subject having a urological

XX disorder.

Claim 1; SEQ ID NO 111; 542bp; English.

The present invention describes the use of polypeptides related to urological disorders for identifying a compound capable of treating a urological disorder, identifying a subject having a urological disorder, or treating a subject having a urological disorder. Also described: (1) a method for identifying a compound capable of treating a urological disorder; (2) a method for identifying a subject having a urological disorder; and (3) a method for treating a subject having a urological disorder. The compound has uropathic and cytostatic activities. The polypeptides related to urological disorders are useful for identifying a compound capable of treating a urological disorder, identifying a subject

CC having a urological disorder, or treating a subject having a urological disorder. Disorders include urinary incontinence and benign prostatic hyperplasia. The present sequence encodes a human urological disorder related protein, which is used in the exemplification of the present invention.

CC Sequence 3931 BP; 1204 A; 725 C; 699 G; 1303 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 13; Length 3931;
Best Local Similarity 54.3%; Pred. No. 1e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGACGCGGTGACATGGAATGCTAAGATTGTGTGCAAGCATCTGTG 448
DB 710 ATTATATATGAGCCAGGGGTGATGAGGCCAAAAGACTGTAACCCAGATGGAAG 769
QY 449 TTGCTATTATGTCCTATAGATGCTGTAAGACACTTTGAGATTAATCTCCAAAT 508
DB 770 TGGCAGTTCATCTCTTCCGTAATCGCATGAACATCTTCCAAATTTTCTTACATC 829
QY 509 TGCACGTGTCCTGCGCAACCAATTTGACATGCAATTTTTCATTTGTGAGCAAGTGG 568
DB 830 TGAATTCATGCTCCAGACAGACGCGGTGAATTTGCGTTTATGTCATTGAACAGACTG 889
QY 569 CGAATCAGACGTTTAATCGCGGAACTAATGAACGTTGATACGACGTACAGCC 628
DB 890 GCACACACCTTTTAACGTCGATGCTTTTCAATGTGGCTTCAAGAGGCGATGAAG 949
QY 629 TCTAACCATGCGAGTCTTATCTTATATGTCATTTTACCTCCGAAAGATGACCTGA 688
DB 950 ACAGTGTCTGGGACTGTATATCTTCCACATGTGATCATCTACGAAAATGACCGGA 1009
QY 689 ACCTGTACACGTGTCATTAACACGATGATGATGATGATGATGATGATGATGATG 748
DB 1010 ACTATTCGATGTGGAAGAAATGACAGTCAATTTTGTGCAAACTGGATTAATCACT 1069
QY 749 ATAACTTCATATTCGCGATCTTTCGCGGAACTGATGATGATGATGATGATGATG 808
DB 1070 ATATTTCTTCATATTAAGAAATTTTGTGTGTGATGATGATGATGATGATGATG 1129
QY 809 AGAAATTCATGATTTTCCATGATTTTGGGTTTGGGTCGAGAGAGACGATTTGG 868
DB 1130 GAAAGATCAATGTTTCTTAATGCTTCTGGGGATGGGAGGAAAGATGATGATCT 1189
QY 869 CGACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 928
DB 1190 GGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1249
QY 929 ATAAATGATTAAGCACTGACGAGACGAGAACTGATTAATTAATGCTGACAAA 988
DB 1250 ACAAGTCAATTCCTCA-TCAACATAGAGGTGAAGTCCAGTT--TTAGGACGGTAAAT 1306
QY 989 TAATGGCCCAAGACGAGCGGATGACAGTGAAGCTTGAAGATTAAGC 1048
DB 1307 TACTTAAGGTATTCGACAGAGCTGATGATGATGATGATGATGATGATGATG 1366
QY 1049 TCGTAATCTGG 1060
DB 1367 CAAAATTAATCTGG 1378

RESULT 11

ID ADR44887 standard; DNA; 3931 BP.

AC ADR44887;

DT 18-NOV-2004 (first entry)

DE Polypeptide 21165 DNA sequence.

XX de; gene; analgesic; gene therapy; pain; painful disorder.

OS Homo sapiens.

PN WO2004071411-A2.

XX 26-ANG-2004.

PF 30-JAN-2004; 2004MO-US002851.

XX 04-FEB-2003; 2003US-0444781P.

PR 05-MAR-2003; 2003US-0452291P.

PR 13-MAR-2003; 2003US-0454540P.

PR 16-JUN-2003; 2003US-0478805P.

PR 30-JUL-2003; 2003US-0491048P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Silvio-Santiago I, Karicheti V, Eliasof SD;

XX WPI; 2004-625760/60.

DR P-PSDB; ADR44888.

PS Disclosure; SEQ ID NO 5; 233bp; English.

CC The invention relates to a method of identifying a compound capable of treating pain or painful disorder by combining a compound to be tested with an e.g., 16386, 15402, 21165, 1423, 636, or 32620 polypeptide or with a host cell expressing the polypeptide and detecting the binding of the CC test compound to the polypeptide to identify a compound that binds to the polypeptide. The method is useful in identifying a compound capable of treating pain or painful disorder. This sequence corresponds to the DNA encoding polypeptide 21165 - a lactosyl ceramide synthase protein.

CC Sequence 3931 BP; 1204 A; 725 C; 699 G; 1303 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 13; Length 3931;
Best Local Similarity 54.3%; Pred. No. 1e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGACGCGGTGACATGGAATGCTAAGATTGTGTGCAAGCATCTGTG 448
DB 710 ATTATATATGAGCCAGGGGTGATGAGGCCAAAAGACTGTAACCCAGATGGAAG 769
QY 449 TTGCTATTATGTCCTATAGATGCTGTAAGACACTTTGAGATTAATCTCCAAAT 508
DB 770 TGGCAGTTCATCTCTTCCGTAATCGCATGAACATCTTCCAAATTTTCTTACATC 829
QY 509 TGCACGTGTCCTGCGCAACCAATTTGACATGCAATTTTTCATTTGTGAGCAAGTGG 568
DB 830 TGAATTCATGCTCCAGACAGACGCGGTGAATTTGCGTTTATGTCATTGAACAGACTG 889
QY 569 CGAATCAGACGTTTAATCGCGGAACTAATGAACGTTGATACGACGTACAGCC 628
DB 890 GCACACACCTTTTAACGTCGATGCTTTTCAATGTGGCTTCAAGAGGCGATGAAG 949
QY 629 TCTAACCATGCGAGTCTTATCTTATATGTCATTTTACCTCCGAAAGATGACCTGA 688
DB 950 ACAGTGTCTGGGACTGTATATCTTCCACATGTGATGATGATGATGATGATGATG 1009
QY 689 ACCTGTACACGTGTCATTAACACGATGATGATGATGATGATGATGATGATGATG 748
DB 1010 ACTATTCGATGTGGAAGAAATGACAGTCAATTTTGTGCAAACTGGATTAATCACT 1069
QY 749 ATAACTTCATATTCGCGATCTTTCGCGGAACTGATGATGATGATGATGATGATG 808
DB 1070 ATATTTCTTCATATTAAGAAATTTTGTGTGTGATGATGATGATGATGATGATG 1129
QY 809 AGAAATTCATGATTTTCCATGATTTTGGGTTTGGGTCGAGAGAGACGATTTGG 868

Db 1130 GAAAGATCAATGTTTCTTCTTAATGCTTCTTGCGGATGGGAGAGAGATGATGACCTTT 1189
QY 869 CGACGAGAACATCGATGCTGCTGAGCTGAAATGTTCAAGATATCGACACAAATTCGACGAT 928
Db 1190 GGAACAGAGTTCACTATGCTGATGATTAATGTAACCAAGACGAGAGCTTAGGAAAT 1249
QY 939 ATAAATGATTAAGCACTGACGAGAGCGACGATTCAGTTAATTAATGCCCTACAAA 988
Db 1250 ACAAGTCATTTCTCTCA-TCAACATAGAGGTGAAGTCACTT--TTTAGACGCTAATAAT 1306
QY 989 TTAATGGCCAAACGAGGCCCGATGACACGTCGCTTAAGCAATCTGAATATAGC 1048
Db 1307 TACTAAGGTATTCGAAAGGCGTCAGTACATGATGACGTAACAATTAATATATAGGC 1366
QY 1049 TCGTAATCTGG 1060
Db 1367 CAAAATCTGG 1378

RESULT 12
AAV73011
ID AAV73011 standard; cDNA; 1853 BP.

XX AAV73011;

DT 15-MAR-1999 (first entry)

DE Human adult brain secreted protein ges1_1 cDNA.

XX Secreted protein; protein factor; human; ges1_1; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 154..975

FT /note= "a
/note= "polynucleotides comprising nucleotides 154-972
and 1-341 are also claimed"

XX PN W09842741-A2.

PD 01-OCT-1998.

XX PF 25-MAR-1998; 98WO-US005972.

XX PR 25-MAR-1997; 97US-00825145.

XX PR 24-MAR-1998; 98US-0046881.

XX PA (GENY) GENETICS INST INC.

PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

DR Spaulding V, Agoetino MJ;

XX WPI; 1999-045135/04.

XX P-PSDB; AAW82006.

XX PT New polynucleotides encoding secreted human proteins - are derived from

XX PT human foetal brain, adult testes, adult brain, or placenta cDNA

XX PT antiinflammatory agents.

XX PS Claim 35a; Page 85-86; 113pp; English.

XX This cDNA sequence codes for novel human secreted protein ges1_1 (see

XX AAW82006). The cDNA was isolated from a human adult brain cDNA library

XX using methods which are selective for cDNAs encoding secreted proteins,

XX or was identified as encoding a secreted or transmembrane protein on the

XX basis of computer analysis of the amino acid sequence of the encoded

XX protein. The nucleotide sequence shows homology to some known database

XX sequences. ges1_1 is a full-length clone and can be obtained from the

XX deposit clone ATCC 98371 as a 1850 bp EcoRI/KcoI fragment. The invention

XX provides polynucleotides (see AAV7300-12), derived from human foetal

XX brain, adult testis, adult brain and adult placenta cDNA libraries, that

CC encode novel secreted proteins (see AAW81997-07). These can be isolated
CC from composite clone ATCC 98371 using specific probes (see AAV73013-22).
CC The polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals (no supporting data
CC are given). Suggested activities include nutritional, immune stimulating
CC (e.g. as vaccines) or suppressing, haematopoietic, tissue
CC growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour invasion
CC suppressor, and tumour inhibition activities. The polynucleotides are
CC also stated to be useful for gene therapy

XX Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;

Query Match 13.6%; Score 157.2; DB 2; Length 1853;
Best Local Similarity 54.4%; Pred. No. 4,2e-40;
Matches 360; Conservative 0; Mismatches 298; Indels 4; Gaps 2;

QY 421 CCTAAGATTTGTTGCAAGGATGCTGTTCTATTTATTTGTCCTTATAGAGATGCTGA 480
Db 499 CCTGAGGTTGTGAGCCCGCTCCGAAACAGCATATTTGCTCATGTCGCCGGAG 558
QY 481 GCACTTTGAGATTAATGCTCCCAATTTGCACTGCTGTCGCCAAACAAATTTGAC 540
Db 559 CACCACTGCGCTGCTGCTCTACCACTGCACTCCCTTCTGCAAGCGCAGAGCTTGT 618
QY 541 TATGCAATTTTCATTTGTTGAGGAGGAGTGGGAGATCAAGCTTTAATCGGGGAACTAATG 600
Db 619 TATGCAATTTTCATTTGTTGAGGAGGAGTGGGAGATCAAGCTTTAATCGGGGAACTAATG 678
QY 601 AACGTTGATACGACGATGATCAAGCTTCACTCCATGAGTGTCTTATCTTTATGAT 660
Db 679 AACGTTGAGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 738
QY 661 GTGATTTACTGCTCCGAAAGATGACCTGTAACCTGTACAGTGT--CAATTCACAGCT 717
Db 739 GTGAGCTCTTGGCCGAAAGATGACCAATCTGTATGTGTGACCCCGGGGAGCCCGC 798
QY 718 CATATGAGTGTAGGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 777
Db 799 CATGTTGCCGTTGCTATGAAACAAGTTGATGATGATGATGATGATGATGATGATGAT 858
QY 778 GGAATCAGTGCATTAACAAAGATGATGATGATGATGATGATGATGATGATGATGAT 837
Db 859 AGAGTCTGAGCACTTACTCTGACCACTGATGATGATGATGATGATGATGATGATGAT 918
QY 838 TGGGTTTGGGCGGAG 897
Db 919 TGGGCTGCGGCGGAG 977
QY 898 GTTCAAGATATCCGACCAATTTGCAAGATTAATTAATTAATTAATTAATTAATTAAT 957
Db 978 ATCTCTGAGCCGCCCACTGTGAGACATTAATTAATTAATTAATTAATTAATTAAT 1037
QY 958 ACGATTCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1017
Db 1038 GGCATATGAGAAATCCCAAGATTTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097
QY 1018 CGTGACGCTTGAAGATTTGAAGTATTAATTTGGAATTTGAAGTATTAATTTGAT 1077
Db 1098 CAAGATGGAGATCACTGACATACAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
QY 1078 AC 1079
Db 1158 AC 1159

RESULT 13
ABQ92036
ID ABQ92036 standard; cDNA; 1853 BP.

XX ABQ92036;

XX

DT 04-OCT-2002 (first entry)
 XX Human polynucleotide SEQ ID NO 33.
 DE
 XX Human; cytostatic; antirheumatic; antiarthritic; vulnery; analgesic;
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KW anticancer; fungicide; antidiabetic; antiaesthetic; antiallergic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 PN US2002065394-A1.
 PD 30-MAY-2002.
 XX
 XX 22-DEC-2000; 2000US-00745763.
 XX
 XX 18-MAR-1998; 98US-00040963.
 XX
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (SPAU/) SPAULDING V.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 DR P-PSDB; ABP61819.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.
 XX
 PS Claim 11; Page 144-145; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,

CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention

XX Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;

Query Match 13.6%; Score 157.2; DB 6; Length 1853;
 Best Local Similarity 54.4%; Pred. No. 4.2e-40;
 Matches 360; Conservative 0; Mismatches 298; Indels 4; Gaps 2;

421 CCTAAGATTGTTGTTCCAAAGCATGCTGTTCTATTATTTGCCCCATAGAGATCGTGA 480
 DB 499 CCGGAGGTTGGAGGCCCGCTCCGAAACAGCATATTGTGCTCATGTGCGCCGGAG 558
 QY 481 GCACATTGAGATAATATGCTCCACAAATTTGCACTGCTGCTGCCAAACAAATTTGAC 540
 DB 559 CACCACTTGCGCGCTGCTGCTCTTACACCTGCAACCTTTCTTGACAGGCCAGCTTGCT 618
 QY 541 TATGCAATTTTCAATTTGAGAGCAAGTGGCAATTCAGACGTTTAAATGCGGGAATTAATG 600
 DB 619 TATGGCATCATGTATCATCCACAGGCTGGAAATGGAACATTAAACAGGGCAAACTGTTG 678
 QY 601 AACGTTGATPACAGCTAGCATGACGCTCTTACCCATGGCAGTCTTCACTTATGAT 660
 DB 679 AACGTTGGGGTGGAGAGGCCCTGCGTGAATGAAGATGGAGATGCTGTTTGGACGAT 738
 QY 661 GTGCAATTTACGCGCGAGATGACCGTAACTGTACACGCTG---CCAAATTCACCACT 717
 DB 739 GTGGACCTCTTGCCAGAAATGACCAACATCTGTATGTGTGACCCCGGGGACCCCGC 798
 QY 718 CATATGATGTATGCGATCGATCAATTAATTAATTTCAATATTTGGCGATCTTGGC 777
 DB 799 CATGTTGCCGTTGCTATGAACAAATTTGGATAGAGCTCCGTAACCCAGTACTCTTGG 858
 QY 778 GGAATAGTGCATCAACAAAGTCACTGGAAGAAATCAATGATTTTGAATGATTTT 837
 DB 859 AGAGTCTCAGCACTTACTCTCTGACGATACCTGAAGATGAATGGCTTCCCAATGATAC 918
 QY 838 TGGGGTTGGGGCGGAGAGCGACGATTTGGCGACAGAACTCATGCTGCTGACTGAAA 897
 DB 919 TGGGGCTGGGGTGTGATGATGACGACTT-GCTACAGGGTGGCGCTGGGATGAAAG 977
 QY 898 GTTTCAGATATTCGACACAAATTTGCAGATATTAATTAAGACTTGAACGAGCG 957
 DB 978 ATCTCTGGCCCCCATATCTGTAGGACACTATAATGATGGTGAAGACCGAGAGATAAG 1037
 QY 958 ACGAATTCAGTTAATTAATGCCCTTACAAATTAATGGCCAAACGAAGCGCGATGGACA 1017
 DB 1038 GGCMAATGAGAAATATCCCAAGATTGACCTCTGTCGTAACCGAATTCCTTGACG 1097
 QY 1018 CGTAGCGGCTTACGATCTGAATTAAGCTGTAATCTGGAATTTGAAGCTCTCTAC 1077
 DB 1098 CAAGATGGAGTGAATCATCATGATACAGTTCGCTGAGAGCTGGGGCTCTTTAT 1157
 QY 1078 AC 1079
 DB 1158 AC 1159

RESULT 14
 ADQ31358
 ID ADQ31358 standard; DNA; 1035 BP.

AC ADQ31358;

DT 07-OCT-2004 (first entry)

DE Human transferase coding sequence, SEQ ID 1.

KW Human; transferase; enzyme;

KW N-acetylglucosamine-6-O sulfate-specific beta 1;
 KW 4-galactosyl transferase; beta 1; 4-galactosylation; gene; ds.

```
XX OS Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..1035
XX FT /tag= a
XX FT /product= "Transferase"
XX PN MO2004058967-A1.
XX PD 15-JUL-2004.
XX PF 26-DEC-2003; 2003WO-JP016948.
XX PR 26-DEC-2002; 2002JP-00378727.
XX (NISC-) JAPAN SCI & TECHNOLOGY AGENCY.
XX Yamashita K, Seko A;
XX WPI; 2004-525884/50.
XX P-PSDB; ADQ31359.
XX Beta 1,4-galactosylating glycan having N-acetylglucosamine-6-O sulfate
XX structure, useful in biosynthesis of sphingosine, involves reacting
XX PT glycan with N-acetylglucosamine 6-O sulfate-specific beta 1,4-galactosyl
XX transferase.
XX PS Claim 3, SEQ ID NO 1, 67pp, Japanese.
XX CC The present invention relates to a human transferase (ADQ31359) with N-
XX CC acetylglucosamine-6-O sulfate-specific beta 1,4-galactosyl transferase
XX CC activity. The transferase is useful for beta 1,4-galactosylating a glycan
XX CC (I) having N-acetylglucosamine-6-O sulfate structure to obtain
XX CC bioynthetic substances. The present sequence is the coding sequence for
XX CC the transferase of the invention.
XX SQ Sequence 1035 BP; 291 A; 229 C; 257 G; 258 T; 0 U; 0 Other;

Query Match 13.5%; Score 155.4; DB 12; Length 1035;
Best Local Similarity 51.8%; Pred. No. 1,2e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTTCAAGAGCATCGTGTGCTATTATGTCCTATAGAGATCGTGA 480
DB 343 CCTCAGGATGTTAAAGCTTTACAGAGGTCGCCATCTGTTCCACCGAAGAGAG 402
QY 481 GGACATTTGAGATTAATGCTCCCAATTTGCACTCGTCTGCCAAGCAATTTGAC 540
DB 403 AAACACCTGATGTAACCTGCTGGAACATCTGCATCCCTTCGTGAGAGGAGAGCTGGAT 462
QY 541 TATGCAATTTTCAATTTGAGCAAGAGGGAATGACAGCTTTAATGCGGAACTTAAG 600
DB 463 TATGGCATCTAGCTCATTCACAGGCTGAAGTAAATTTAATGACCAAACTCTTG 522
QY 601 AACGTTGATGAGAGCTAGCATCAGCGCTTACCCATGGAGAGCTTCACTTTCAATAT 660
DB 523 AATGTCGCTATCTAGAAAGCCCTCAAGAGAAATTTGGAATGCTTTTATTTCCAGAT 582
QY 661 GTGCAATTTTACGCCCCAGAGATGACCGTAACTGTACAGCTGTCCAAATTCACACGCTCAT 720
DB 583 GTGAGACCTGTACCGGAGATGACTTTAACTTTAAGAGTGAAGAGATCCCAAGCAT 642
QY 721 ATGAGGTGAGCGATCATTAATTTCAATTTAACTTCCATTTGCGCGATCTTGGCGGA 780
DB 643 CTGGTGTGGGCGAGAAAGCACTGGGTACAGGTTACGTTACAGTGAATATTTTGGGGGT 702
QY 781 ATCAGTGCACTAAACAAGATGACCTGAAGAAATCAATGATATTTTGAATGATTTTGG 840
DB 703 GTTACTGCTTAAAGAGAGAGCTTTTCAAGGTGAATGATTTCTTAACTAATCTAG 762
QY 841 GGTGGGGCGAGAGAGACGATTTTGGCAGAGAAATCGATGCTGAGCTGAAGATT 900
DB 841 GGTGGGGCGAGAGAGACGATTTTGGCAGAGAAATCGATGCTGAGCTGAAGATT 900
```

```
DB 763 CGATGGGAGGCGAAGACAGTACCTCAGACTCAGGGTTGACTCCAAAGATTAATTT 822
QY 901 TCAAGATATCCGACACAAATTCAGATTAATTAATTAATTAATTAATTAATTAATTA 960
DB 823 TCCCGCCCCCTGCTCGAAGTGGTAAATTAATTAATTAATTAATTAATTAATTAATTA 882
QY 961 AATCCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
DB 883 AATGAGGTGAACGCAAGAGATGAAGCTTTACACCAAGTGTCAAGAGCTGAGAGACA 942
QY 1021 GACGCGCTTAGCAATCTGAATTAAGCTGTAATTTGAAATTTGAAGCTCTTCAACT 1080
DB 943 GATGGGTGAGATGATGTTCTTAAATTAATTAATTAATTAATTAATTAATTAATTA 1002
QY 1081 CGAGCCGTGCTGATTT 1097.
DB 1003 AACATCAAGTGAATTT 1019

RESULT 15
AAZ65035
ID AAZ65035 standard; cDNA; 1964 BP.
XX AAZ65035;
AC 02-JUN-1998; 98US-0087607P.
XX 02-JUN-1998; 98US-0087609P.
XX 02-JUN-1998; 98US-0087759P.
XX 03-JUN-1998; 98US-0087827P.
XX 04-JUN-1998; 98US-0088025P.
XX 04-JUN-1998; 98US-0088028P.
XX 04-JUN-1998; 98US-0088029P.
XX 04-JUN-1998; 98US-0088030P.
XX 04-JUN-1998; 98US-0088033P.
XX 04-JUN-1998; 98US-0088326P.
XX 05-JUN-1998; 98US-0088167P.
XX 05-JUN-1998; 98US-0088202P.
XX 05-JUN-1998; 98US-0088212P.
XX 05-JUN-1998; 98US-0088217P.
XX 05-JUN-1998; 98US-0088655P.
XX 09-JUN-1998; 98US-0088722P.
XX 10-JUN-1998; 98US-0088730P.
XX 10-JUN-1998; 98US-0088734P.
XX 10-JUN-1998; 98US-0088738P.
XX 10-JUN-1998; 98US-0088740P.
XX 10-JUN-1998; 98US-0088741P.
XX 10-JUN-1998; 98US-0088742P.
XX 10-JUN-1998; 98US-0088810P.
XX 10-JUN-1998; 98US-0088811P.
XX 10-JUN-1998; 98US-0088824P.
XX 10-JUN-1998; 98US-0088825P.
XX 10-JUN-1998; 98US-0088826P.
XX 11-JUN-1998; 98US-0088859P.
XX 11-JUN-1998; 98US-0088861P.
XX 11-JUN-1998; 98US-0088863P.
XX 11-JUN-1998; 98US-0088876P.
XX 12-JUN-1998; 98US-0089090P.
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 18:40:30 ; Search time 5190 Seconds
(without alignments)
10385.110 Million cell updates/sec

Title: US-10-661-430-2
Perfect score: 1152
Sequence: 1 atgcttcctgcatttgcgc.....acttccaacgtttttag 1152

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*
10: gb_esc10:*
11: gb_esc11:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	672.4	58.4	860	7	CV125668 OSTR30043
2	643	55.8	866	7	CV125669 OSTR30043
3	178.2	15.5	687	5	BM054310 BM054310
4	177.8	15.4	679	3	BJ061791 BJ061791
5	177.2	15.4	622	5	BM347970 BM347970
6	167.6	14.5	796	6	CX757659 AGENCOURT
7	167.6	14.5	879	6	CD514077 AGENCOURT
8	167.6	14.5	955	5	BQ957788 AGENCOURT
9	167.6	14.5	1196	11	DD035706 Homo sapi
10	167.6	14.5	1590	4	CR614082 full-length
11	167.6	14.5	1860	4	CR612341 full-length
12	167.6	14.5	1870	4	CR607487 full-length
13	167.6	14.5	1876	4	CR592876 full-length
14	167.6	14.5	1876	4	CR594939 full-length
15	165	14.3	673	4	AJ740626 AJ740626
16	164.4	14.3	1979	4	CR925982 Pongo pyg
17	163.6	14.2	402	1	AI987157 rs22603.Y
18	163.4	14.2	715	7	CD081993 MA3-9999U
19	163.4	14.2	715	7	CN555132 tae13a12.
20	163.2	14.2	612	7	CX546970 BwK20-000
21	162.4	14.1	667	5	BM052849 BM052849
22	161.2	14.0	893	8	CX362564 JGI_XZT42

23	161	14.0	661	3	BU617372 BU617372
24	160	13.9	588	5	BU029426 BU029426
25	160	13.9	684	5	BM437339 BM437339
26	157.6	13.7	970	1	AL558425 AL558425
27	157.2	13.6	844	1	AL866706 AL866706
28	155.4	13.5	1035	10	AY410571 Homo sapi
29	155.4	13.5	2203	4	CR749555 Homo sapi
30	155	13.5	561	5	BM044315 BM044315
31	155	13.5	803	8	CX744001 JGI ANBT1
32	154.4	13.4	822	8	DN948717 AGENCOURT
33	153.4	13.3	737	8	CX744145 JGI ANBT1
34	152.8	13.3	774	7	CK026161 AGENCOURT
35	152.4	13.2	1149	10	AY412728 Homo sapi
36	152.2	13.2	907	5	BM488994 BX488994
37	151.6	13.2	690	7	CN280919 170004554
38	151.4	13.1	791	8	CX780748 AGENCOURT
39	151.2	13.1	831	5	BQ918131 AGENCOURT
40	150	13.0	759	8	CX457513 JGI_XZG54
41	150	13.0	1052	10	AY417315 Homo sapi
42	150	13.0	1368	4	CN50GMSO Tetraodon
43	149	12.9	1009	10	AY410572 Pan trogl
44	148.8	12.9	1738	4	BC004095 Mus muscu
45	148.8	12.9	2369	4	BC029303 Mus muscu

ALIGNMENTS

RESULT 1
CV125668 860 bp mRNA linear EST 31-AUG-2004
LOCUS OSTR30043C08 PDONR201 Entry vector Caenorhabditis elegans cDNA,
DEFINITION mRNA sequence.
ACCESSION CV125668.1 GI:51716887
VERSION EST.
KEYWORDS Caenorhabditis elegans
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 860)
Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., Li, N., Sequerra, R.,
Bosak, S., Doucette-Stamm, L., Vandenhaute, J., Hill, D.E. and Vidal, M.
C. elegans ORFeome Version 3.1: Increasing the coverage of ORFeome
resources with improved gene predictions
Genome Res. (ORFeome issue) (2004) In press
JOURNAL Contact: Philippe Lamesch and Tong Hao
COMMENT Marc Vidal Lab
DFCI
44, Binney Street, Boston, MA 02115, USA
Tel: 6176323910
Fax: 6176325739
Email: philippe_lamesch@dfci.harvard.edu
PCR Primers
FORWARD: TGGCTTTCGTCATTGCGCAGT
BACKWARD: TAAACACGCTGGGAAGTCCTT.

FEATURES

source
1..860
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="male and hermaphrodite"
/tissue_type="all tissues"
/cell_type="all cells"
/dev_stage="all stages"
/clone_lib="PDONR201 Entry vector"

ORIGIN

Query Match 58.4%; Score 672.4; DB 7; Length 860;
Best Local Similarity 97.9%; Pred. No. 1.1e-202;
Matches 703; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

DEFINITION BM054310 Nori Satoh unpublished cDNA library, blood cells Ciona intestinalis cDNA clone c1db080a21 5', mRNA sequence.

ACCESSION BM054310

VERSION BM054310.1 GI:24155006

KEYWORDS EST

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.

AUTHORS Satou, Y., Satake, M., Azumi, K., Nonaka, M., Shin-i, T., Kohara, Y. and Satoh, N.

TITLE Expressed genes in Ciona intestinalis (2002)

JOURNAL Unpublished (2002)

COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source location/Qualifiers

1..687

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="c1db080a21"

/tissue_type="blood cells"

/clone_11b="Nori Satoh unpublished cDNA library, blood cells"

ORIGIN

Query Match 15.5%; Score 178.2; DB 5; Length 687;
Best Local Similarity 55.1%; Pred. No. 7.2e-45;
Matches 348; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

448 GTTGCTATTATTTGTCCTTATAGATCGTGAAGACATTTGAGATATGCTCCACAT 507
|||||
30 GTTGCAATGTTATTTCTTATCGGGACAGAGAAGACATTTGCGGATTTCTCGAGTAC 89
|||||
508 TTGCACTGCTGCTGCGCCAAACAACATTTGACTATGCAATTTTCAATTTTGAGAGCAAGTG 567
|||||
90 ATGCAATCCCAACCCCTCAACAGTCAACACTTGATAGCCATCTATGTTGTTATCAGGCT 149
|||||
568 GCGATTCAGACGTTTAAATGCGGGAAACTAATGAAGCTTGATACGAGCTACATCAGCGC 627
|||||
150 GGTACTGGTAAATTTATGTCGCCAAGCTGATGAATGTTGGTTATGCTGATCAATCAAA 209
|||||
628 CTCTACCCATGCGAGTGTTCATCTTTCATGATGTCGATTTTACTGCCCGAAGATGACGT 687
|||||
210 GATCATGATTTTCATGTTTTCCTTCCATGATGTCGATTTGCTGCTTGAGATGACAAG 269
|||||
688 AACCTGTACACGTGTCATTCACCAACGTCATATGATGTAAGCATGATGAATTCAT 747
|||||
270 AGTATATCTCTTGGCCCCCTCAAGTCTCGCACCTGTCAGCTGGGGTGAACAAGTTTAA 329
|||||
748 TATAAATCTCCATTTTCGCGCATCTTCGCGGGAATTCAGTCACTAAACAAATACCTG 807
|||||
330 TACCAACTGCCCTACTCTGCCATATTTGGTGGCTGAGCTGACCTGAACAAAGAACATTC 389
|||||
808 AAGAAATCAATGATTTTTCGATGATTTTTCGGGTGGGGCGAGAGACACGATTTG 867
|||||
390 CAAAAGTGAAGGTTTACTCTTAATTCATCTGGGGTGGGGTGGAGAGACACGATTTG 449
|||||
868 GCGAGAGAACATCGATGCGTGAAGTGAAGTTTCAAGATATCGACACAAATTCGACGA 927
|||||
450 TTTAATCCGTCAGGTTTCTGCGATGATATTTATTCGTTACCTATGATATCTCAAG 509
|||||
928 TATAAATGATTAAGCACTCGACGAGAGAGAAATCCGTTAATTAATGCCGCTACAA 987
|||||
510 TATAAGATATTAATCTCAAGAGAGAGAAAGGAAACGAAACCAACCAAGGCTTCGAC 569
|||||
988 ATTAATGGGCAACGAGCGCGATGACGTCGAGCGGCTTAAGCAATCTGAAGTAAAG 1047

Db 570 CAAATACCGACGACTTAAGACACGATGCGCAACGATGCTTAACACATTAATACAG 629
|||||
Qy 1048 CTCGTAATCTGGAATTTGAAGCTCTCTACAC 1079
|||||
Db 630 GTCTATCAAGACAGAGAACAGACTTTATAC 661
|||||

RESULT 4

LOCUS B061791 679 bp mRNA linear EST 29-SEP-2003

DEFINITION B061791 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone X1074b05 5', mRNA sequence.

ACCESSION B061791

VERSION B061791.1 GI:17423856

KEYWORDS EST

SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus, Xenopus.

AUTHORS Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.

TITLE Expressed genes in X. laevis embryo

JOURNAL Unpublished (2001)

COMMENT Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gens.nig.ac.jp
The information of this clone is available through the following URL.
http://xenopus.nibb.ac.jp.

FEATURES

source location/Qualifiers

1..679

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="X1074b05"

/tissue_type="whole embryo"

/dev_stage="stage 25"

/clone_11b="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN

Query Match 15.4%; Score 177.8; DB 3; Length 679;
Best Local Similarity 53.8%; Pred. No. 9.6e-45;
Matches 365; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

418 ATGCTAAGATGTTGTTGCAAGCATGTTGCTATTTATTTGTCCTTATAGATGCT 477
|||||
2 AGGCCCGACGATTTGATGATGCCAGCAGNAAGTTGCAATTAATCCATTCGTTATGCA 61
|||||
478 GAAGACATTTGAAGTAATGTCACATTTGCACTGCTGCTGCCAAACAATTTG 537
|||||
62 GAAGCTCATCTCAAGACATGCTGTAATCAAGCACCGCTTCACGCGGACGACGA 121
|||||
538 GACTATGCAATTTTCATTTGAGAGCAAGTGGCAATCGAGCTTAATTCGCGGAACTA 597
|||||
122 GATTATGCTGATATATGCTGTGGAACAGATGGAAGACAGATTTTGAACGCGCAAGCTG 181
|||||
598 ATGAACGTTGATACGACGATGATCAGCGCTTACCAATGAGCACTGCTTATCTTCA 657
|||||
182 ATGAATGTTGTTATGCGACAGCATCAAAAGACTATGATTTATGCTGTTTATATTCAG 241
|||||
658 GATGCGATTTTACGCGCGGAGATGACCGTAACTGACGTCAGTTCGAATTCACAGCT 717
|||||
242 GATGCGACATTTATCCCTATGATGAGAGAAATCTCAAGGCTCTCAAAATTCGACA 301
|||||
718 CATATGAGTGAAGCATGATTAATTCATTAATTAATTCATATTCGCGCATCTTCGCG 777

Db 302 CATATGGCAAAATTCGTGGACAACTTTAACTTTAACTTACCTTACATGACATCTTTGGA 361
Qy 778 GGAATGACGTGACCTAAACAAAGATCACTCGAAGAAAATCAATGATTTTGCATGATTTT 837
Db 362 GGAAGTTGGGCAATTCACAAAAGAACCTTATATCAAAAGTGAATGATTTCTCAATGTTTC 421
Qy 838 TGGGGTTGGGGGGAAGAGACGACGATTTTGGGACGAGAACATCGATTTGGCTGGACTGAAA 897
Db 422 TGGGGGTGGGAGAGAGATGATGAACTGTTCCAAAAGAGTGTGCTGTGGAGCTGAAA 481
Qy 898 GTTTCAGATATCCGACACAAATTTGACATATTAATTAATTAAGATCTGACGAGAGCG 957
Db 482 GTTGGACGCTGCAGACCAAACTATGGCAATCCAGATGATTTCTCATTAAGAGATCCG 541
Qy 958 AGCAATCCGATTATTAATGCGCTACAAATATATGGCCCAAGCAAGCGCGATGACA 1017
Db 542 GGTAAATGAAGAAGAAAGAGCTTCTCTATTAAGAAAGCCCTATTAACGCTGCAAT 601
Qy 1018 CGTGAAGGCTTACGACATCTGAAGTAAAGCTGTAATCTGGAATTTGAAGCCTCTAC 1077
Db 602 CAGGATGAGACTGATTCCTGGAATTAACCATATATCATATTAATGATCTAAACTTTAT 661
Qy 1078 ACTCGAGCCGCTGCTGAT 1095
Db 662 ACAAAGTCACAGTGAT 679

RESULT 5
BM347970 622 bp mRNA linear EST 27-MAY-2004
LOCUS BM347970 Yutaka Satou unpublished cDNA library, embryo whole animal
DEFINITION Clona intestinalis cDNA clone ciem838b10 5', mRNA sequence.
ACCESSION BM347970 GI:47759771
VERSION BM347970.1 GI:47759771
KEYWORDS Clona intestinalis
SOURCE Clona intestinalis
ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cloniidae; Clona.

REFERENCE 1 (bases 1 to 622)
AUTHORS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Clona intestinalis (2004)
JOURNAL Unpublished (2004)
COMMENT Contact: Yutaka Satou
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113
Email: yutaka@sci.kyoto-u.ac.jp.
Location/Qualifiers

FEATURES
source
1..622
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciem838b10"
/tissue_type="whole animal"
/dev_stage="embryo"
/clone_id="Yutaka Satou unpublished cDNA library, embryo
whole animal"

ORIGIN

Query Match 15.4%; Score 177.2; DB 5; Length 622;
Best Local Similarity 55.8%; Pred. No. 1.4e-44;
Matches 338; Conservative 0; Mismatches 268; Indels 0; Gaps 0;
Qy 448 GTTGCTATTATTTGGCCCTATAGATCGTGAAGCACATTTGAGATTAATGCTCCACAAT 507
Db 12 GTTGCCATAGTATTTCTTATCGGACAGAGAAACAATTTGCGATTTTCTCTGAGTAC 71
Qy 508 TTGCACTCGTTGCTCGCCAAACAATTTGACATATCTTCAATTTGAGAGCAAGTG 567

Db 72 ATGATCCCAACCCCTCAACGTCACAACTTATTAAGCCATCTATGTTGTTAATGAGCT 131
Qy 568 GGAATGACAGCTTTAATGCGGGGAAACTATGAACGTTGATGACGATCAGATCAGC 627
Db 132 GGTACTGTTAAATTAATGATCCAGACTGATGAATGTTGGTTATGCTGATCAATCAAA 191
Qy 628 CTCTACCCATGGCAGTCTTCACTTTTCATGATGTCGATTTTAATCTGCCGAGATGACCGT 687
Db 192 GATCATGATTTTCATATGTTTCTTCTTCCATGATGAGACTTGCTGTTGAGATGACAAG 251
Qy 688 AACCTGTACAGCTGTCCAAATTAACACGCTCATTAAGTGTAGCGATGATTAATTCAT 747
Db 252 AGTATATACCTTTCGCCCTCAAGTCCCTGCCACCTGTACGCTGGGGTGAACAAGTTTAA 311
Qy 748 TATTAACCTTCAATTTTGGCGATCTTGGCGGATCAAGTACGACCTAACAAAGATCACTG 807
Db 312 TACCAATGCCCCCTACTCTGCAATATTTGGTGGCGTGACTGACCTGACAAAGAACTTC 371
Qy 808 AAGAAATCAATGATTTTTCGAATGATTTTGGGTTGGGGCGGAGAGACGACGATTTG 867
Db 372 CAAAAGTGAAGGTTTACTCTTAATTCATTCTGGGGTGGGGTGGAGAAAGACGATATG 431
Qy 868 GCGACGAGAACATCGATGCGCTGACGAAAGTTTCAAGTATCCGACAAATTGACGA 927
Db 432 TTTAACCGTGTCAAGTTTCTGGCATGAATATTTATTCCTATGATATCTTCAAG 491
Qy 928 TATTAAGATTAAGACATCGACGAGGAGGAGAAATCCAGTTAATTAATGCGCTCAAA 987
Db 492 TATTAAGATATTAATCTATCAAGAGAGAGAAAGGAAAGCAACCAACCAAGCTTCGAC 551
Qy 988 ATTAATGGGCAACGAAAGCGGATGACACGTCGACGCTTAAGCAATCTGAAGTAAAG 1047
Db 552 CAATACGACGAGCTTAAGACACGATGCAAGAGATGCTTAACACACTTGAATTAAG 611
Qy 1048 CTCGTA 1053
Db 612 GTCGTA 617

RESULT 6
CX757659 796 bp mRNA linear EST 24-JAN-2005
LOCUS CX757659
DEFINITION AGENCOUNT 40973844 NIH-MGC_281 Homo sapiens cDNA clone
IMAGE:7782009 3', mRNA sequence.
ACCESSION CX757659 GI:58054315
VERSION CX757659.1 GI:58054315
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 796)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Meri Firpo
cDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM15939 row: p column: 07
High quality sequence start: 16
High quality sequence stop: 792.
Location/Qualifiers

FEATURES
source
1..796

QY	661	GTGATTTACTGCCCCGAAGATGACCGTAACTGTACACAGTGT-----CGAATTCACCAAGT	717
Db	285	GTGACACTCTTGGCCGAAAAATGACCACATCTGTATGTGTGTGA.CCCCCGGGGACCCCGC	344
QY	718	CATATGATGTAGCGATCGATTAATTTCAATTATATAACTTCCATATATTCGGGCATCTTGGGC	777
Db	345	CATTTCGCCGTTGTATGAACAAGTTTGGATACAGCCCTCCGTTACCCCACTACTTGGGA	404
QY	778	GGATTCAGTGCATTAACAAAAGATCACTGAAAGAAATCAA.TGGATTTTGCATAGATTTT	837
Db	405	GGATCTTCAGAGACTTACCTCTGACCACTGAAAGATGGAATGGCTTCCCAATGAATAC	464
QY	838	TGGGGTTGGGGCCGGAGAGACGACGATTTGGCCGACGAAACATCGATGGCTGCACTGAAA	897
Db	465	TGGGGCTGGGGTGTGTAGAGATGACGACATTGCTACAGGGGTGGCGCTGTGGATGAG	524
QY	898	GTTTCAAGATATCCGACACAAATTGCA.CGATATATAAATGATTAAAGCACTCGACGGAAACGG	957
Db	525	ATCTCTGGGCCCCCA.CATCTGTAGGACACTATATAAGATGTGAAGCACCGAGAGATTAAG	584
QY	958	ACGAATTCAGTTAATTAATATGCCCGCTACAAATAATATGGGCCCAAACGAGCGCCGATGCA	1017
Db	585	GGCAATAGAGAAATTCGCCACAGATTGACCTCCTGTGTCGATACCAAAATTCCTGGACG	644
QY	1018	CGTACGGCCCTAACCAATCTGAAGTATAGCTCGTAATCTGGAATTGAACCTCTAC	1077
Db	645	CAAGATGGAGATGAAC.TCACTGACATCAAGTTGCTGSGCTCGAGAGCTGGGGCTCTTTAT	704
QY	1078	AC 1079	
Db	705	AC 706	

LOCUS	955 bp	mrna	linear	EST 21-AUG-2002
DEFINITION	BO957788			
AGENCY/ACCESSION	AGENCOCRT_8772708	NIH_MGC_18	Homo sapiens	cdna clone IMAGE:6372348
VERSION	BO957788			
KEYWORDS	BO957788.1	GI:22373266		
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo			
REFERENCE	1 (bases 1 to 955)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DTF/Gazdar cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LNCM2547 row: p column: 13 High quality sequence Etop: 617. Location/Qualifiers 1..955			
FEATURES				
source				

FEATURES	SOURCE	Location/Qualifiers
		1..955
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="IMAGE:6372348"
		/tissue_type="large cell carcinoma"
		/lab_host="DH10B (phage-resistant)"
		/clone_idb="N1H_MGC_18"
		/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG (c). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

Query Match	Similarity	Score	DB	Length
Best Local	54.4%	Pred. No. 1.9e-41		
Matches	360	Conservative	0	Mismatches 239, Indels 3, Gaps 1
QY	421	CCTAAGATTGTTGTGGCAAGGCACTCGTGTGCTATTATTTAGTCCCTTATAGATCTGAA	480	
DB	79	CTTCAGATTGTATGAGCCCGCTCCGAAAGCATATTTGCTCATTCGCTGAGCCGGAG	138	
QY	481	GCACATTGGAAATATGCTTCACAAATTTGACCTGTTGCTCGCCAAACAAATTTGAC	540	
DB	139	CACCACTGGCGCTGCTGCTTCACACCTGACCCCTTTGACGCGCAGACACTTGT	198	
QY	541	TATGCAATTTTCATTGTTGAGAGCAAGTGGCCAAATCAGACGTTTATTCGGGAAACTATG	600	
DB	199	TATGGCAATCTATATGTCATCCACAGGCTGGAATGAAACATTTTAAAGGCAAACTGTG	258	
QY	601	AACGTTGGATACAGCTAGCATCAAGCCTCAACCCATGSGCAGTCTTCATCTTCATGAT	660	
DB	259	AACGTTGGGCTGCAAGAGCCCTGCTGATGTAAGTGGGACTGCTGTTCTTGACAGAT	318	
QY	661	GTCGATTTTACGCCCCGAAGATGACCGTAACTGTACGTGT--CCAATTCACCACT	717	
DB	319	GTGAGCTCTTGCCAGAAATGACCAACATCTGTATGTGTGACCCCGGGGACCCGC	378	
QY	718	CATATGAGTGTAGGATGATGATTAATTTCAATATATACTTCCATTTGGCGATCTTGGC	777	
DB	379	CATGTGTCCGTGTCTATGATGAAACAAATTTGGATGACGCTCCGTAACCCCACTTCTGGA	438	
QY	778	GGAATCAGTGCACTAACAAAGATCACCTGTAAGAAATTCATGGATTTTCCATGATTTT	837	
DB	439	GGAATCTAGCATTTACTCTGACCACTGTAAGTGAATGCTCTCCCAATGAATAC	498	
QY	838	TGGGTTGGGGCGGAGAGACGACGATTTGGCGACGAAACATGATGCTGACCTGAA	897	
DB	499	TGGGCTGGGGTGGTGAAGATGACGACATTTCTACAGGGGTGCGCTGGCTGGATGAG	558	
QY	898	GTTTCAAGATATCCGACACAAATTTGACAGATATATAATTAAGTCACTCGACGAAACG	957	
DB	559	ATCTCTCGGCCCCCACTGTGTGGAACATATAGATGTGAAGCAACCGAGAGATAG	618	
QY	958	ACGAATCAGTTAATTAATGCGGCTACAAATAATGSGCCAAAGAGCGCCGATGACA	1017	
DB	619	GGCAATGAGAAATCCCAACAGATTTGACTCTGCTGCTGCTGATCCGAATTTCTGAGC	678	
QY	1018	CGTACGCGCTTAAGCAATCTGAAGTATAGCTCGTAAATTCGGAATGMACTCTCTAC	1077	
DB	679	CAAGATGGAGTGAATCACTGACATACACAGTTGCTGCTGAGAGCTGGGCTCTTTAT	738	
QY	1078	AC 1079		
DB	739	AC 740		

RESULT 9	DQ035706	1196 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ035706				
DEFINITION	Home sapiens B4GALT3 gene, VIRTUAL genomic survey sequence.				
ACCESSION	DQ035706				
VERSION	DQ035706.1				
KEYWORDS	GSS.				
SOURCE	Home sapiens (human)				
ORGANISM	Home sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED
1 (bases 1 to 1196)	Nielsen, R., Bustamante, C., Clark, A.G., Gianowski, S., Sackton, T.B., Hubisz, M.J., Fedel-Alon, A., Tenebaum, D.M., Civello, D., White, T.J., Sniinsky, J.J., Adams, M.D. and Cargill, M.	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees	<i>PLoS Biol.</i> 3 (6), E170 (2005)	15869325
2 (bases 1 to 1196)	Nielsen, R., Bustamante, C., Clark, A.G., Gianowski, S., Sackton, T.B., Hubisz, M.J., Fedel-Alon, A., Tenebaum, D.M., Civello, D., White, T.J., Sniinsky, J.J., Adams, M.D. and Cargill, M.	Direct Substitution		
TITLE		Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT		This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.		

```

FEATURES
  source
    location/Qualifiers
      1..1196
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="1"
      <1..21196
        /gene="B4GALT3"
        /locus_tag="HC4901"

```

ORIGIN

Query Match	14.5%	Score 167.6	DB 11	Length 1196
Best Local Similarity	54.4%	Pred. No. 2.1e-41		
Matches 360	Conservative 0	Mismatches 299	Indels 3	Gaps 1

[illegible]

QY 1018 CGGACGGCCCTAAGCATCTGAAGTATPAAGCTGTAATCGAATGAAGCCCTCTAC 1077
 |||||
 Db 960 CAAGATGGATGAATCACTGACTGACATCAACAGTTGCTGGCTGCAAGCTGGGGCCTCTTAT 1019
 |||||
 QY 1078 AC 1079
 |||||
 Db 1020 AC 1021

RESULT	10
LOCUS	CR614082
DEFINITION	CR614082 1590 bp mRNA linear HTC 21-JUL-2004 full-length cDNA clone C50DU005YK01 of T cells (Jurkat cell line)
ACCESSION	CR614082
VERSION	CR614082.1 GI:50494889
KEYWORDS	HTC; CNSLT_CDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE

AUTHORS Li, W. B., Gruber, C., Jesse, J., and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@liffech.com URL :

REFERENCE

Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL TITLE

COMMENT

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES	Location/Qualifiers
source	1. .1590

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/ciOne="CS0DJ005YK01"
/rtsbse_type="T cells (Jurkat cell line) Cotte
10-normalized"
/plasmid="pCMVSPORT_6"

```

Query Match 14.5%; Score 167.6; DB 4; length 1590;
 Best Local Similarity 54.4%; Pred. No. 2.3e-41;
 Matches 360; Conservative 0; Mismatches 259; Indels 3; Gaps 1;

Qy	421	CCTAAGGATTTGTGTGCAAGGCAATCGTGTGCTATTATTTGTGCCCTTAATGAGATCGTGAA	480
Db	282	CTCGACAGTGTGTAGGCCCCGCTCCGAAACAGCATCTTGTGCTCATCGGCCCGGGAG	341
Qy	481	GCACATTTGGAATTAATGTCTCCACAAATTTGCACGTGTCTCGGCAACAACAATTTGAC	540
Db	342	CACCACTGTGGCCTGTGCTCTTACACCTGCACCCCTTTCTTGACGCCACGACCTTGCT	401
Qy	541	TATGCAATTTTCAATTTGTGAGCAAGTGGGCAATTCAGCGTTTATGCGGGAAACTATG	600
Db	402	TATGGCATCTATGTCTATCCACAGCTGAGAAATGAAATTTAACAGGCAAACTGTG	461
Qy	601	AACGTTGGAATACAGCTAGATCAAGCCTTACCATGGCAGTGTCTCATTTTCATGAT	660
Db	462	AACGTTGGGGTGGAGAGGCGCTGCGATGAAGAAGTGGACCTGCTTCTTTCACAGAT	521
Qy	661	GTGCATTTACTGCCCCGAAGATGACCGTAACTGTACCGTGT--CCAAATTCACCAAGT	717

Db 522 GTGACCTCTTGGCCAGAAAAATGACCAAACTGTATGTGTGATGACCCCGGGAGACCCCG 581
Qy 718 CATATAGGTGAGCGATTCGATTAATCAATTATTAACCTTCATATTCGGGATCTTGCGC 777
Db 582 CATGTGCGCGTGTATGAACAAGTTTGATACAGCCTCCGTAACCCCAAGTACTTCGGA 641
Qy 778 GGAATGAGTGCATCAACAAAGATTCACCTGAGAAAAATCAATGATTTTTCGATGATTTT 837
Db 642 GGAAGTTCAGCACTTACTCTGACCAAGTACCTGAAGATGATGATGCTTCCCAATGATAC 701
Qy 838 TGGGGTTGGGGCGGAGAGAGACGATTTGGCGACGAGAAATCATGCTGCTGACCTGAAA 897
Db 702 TGGGGCTGGGGTGGTGAAGATGACGACATTTGCTACCAAGGTGGCGCTGGGATGAAAG 761
Qy 898 GTTTCAGATATGCGGACACAAATTTGACGATATTAATGATTAAGCACTGACGGAAGCG 957
Db 762 ATCTTCGGCCCCCAGATCTGTAGGACACTATTAAGTGTGAGAGACGAGAGATTAAG 821
Qy 958 ACGAATCCAGTTAATAATGCGCTACCAAAATATGAGCCCAAGAGCGCCGATGAGCA 1017
Db 822 GCGAATGAGGAAATATCCCAAGATTTGACCTCTGCTCGTACCAAGATTTCTGAGAG 881
Qy 1018 CGTACGCGCTTAACGATCTGAAGTATAGCTGCTAATCTGGAATTTGAAGCTCTCTAC 1077
Db 882 CAAGATGGGATGAATCACTGACATACCAAGTTGCTGCTGAGAGCTGGGGCTCTTTAT 941
Qy 1078 AC 1079
Db 942 AC 943

RESULT 11
CR612341 1860 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DK011Y18 of HeLa cells Cot 25-normalized
DEFINITION

CR612341
CR612341.1 GI:50493148
HTC; CDS/UT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1860)

REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1. 1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011Y18"
/issue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 14.5%; Score 167.6; DB 4; Length 1860;

Best Local Similarity 54.4%; Pred. No. 2,4e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

Qy 421 CCTAAGATTTGTTGTCAGAGGATGCTGTGCTATATATGTCCTTATGAGATGCTGAA 480
Db 562 CCGCAGGTGTGAGCCCGCTCCGACAGCAATCATTTGTGCTTATGTCGCCGGAG 621
Qy 481 GCACATTTGAGATATATGCTCCCAATTTGCACTGCTTCTGCGCAACAAATTTGAG 540
Db 622 CACCACTGCGCTGTGCTTACCACTGCAACCCCTTCTTGAGGCGCAGCAGCTTGCT 681
Qy 541 TATGCAATTTTCAATTTGTGAGAGCAATGCGCAATCAAGCTTTATATGCGGAACTAATG 600
Db 682 TATGCAATTTTCAATTTGTGAGAGCAATGCGCAATCAAGCTTTAATCAAGGCAAACTGTG 741
Qy 601 AACGTTGATGAGAGCTAGCATCAAGCTCTTACCCATGAGCTTCACTTTTCATGAT 660
Db 742 AACGTTGAGGTCGAGAGAGCCCTGCTGATGAGAGTGGAGCTGCTGTTTTCACGAT 801
Qy 661 GTGATTTTACTGCGCCGAGAGATGACCGTAACTGTACACGTG---CAATTAACGACGT 717
Db 802 GTGACCTCTTGGCAAGAAATGACCAATCTGTATGTGTGACCCCGGGAGACCCCG 861
Qy 718 CATATAGGTGAGCGATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 777
Db 862 CATGTGCGGTCATGAAACAAGTTTGATACAGCTCCCGTACCCCAAGTACTTCGGA 921
Qy 778 GGAATGAGTGCATCAACAAAGATTCACCTGAGAAATCAATGATTTTTCGATGATTTT 837
Db 922 GAGGTCTCAGCACTTACTCTGACAGTACCTGAGATGAATGATTTTCCCAATGATAC 981
Qy 838 TGGGGTTGGGGCGGAGAGAGACGATTTGGCGACGAGAAATCATGCTGCTGACTGAAA 897
Db 982 TGGGGCTGGGGTGGTGAAGATGACGACATTTGCTACAGAGGTGGCTGCTGGAGTAAAG 1041
Qy 898 GTTTCAGATATTCGACCAAAATTTGACGATATTAATGATTAAGCACTGACGGAAGCG 957
Db 1042 ATCTTCGGCCCCCAGATCTGTAGGACACTATTAATGATGATGAGAGACGAGAGATTAAG 1101
Qy 958 ACGAATCCAGTTAATAATGCGCTACCAAAATATGAGCCCAAGAGCGCCGATGAGCA 1017
Db 1102 GCGAATGAGGAAATATCCCAAGATTTGACCTCTGCTGCTGACCAAGATTTCTGAGAG 1161
Qy 1018 CGTACGCGCTTAACGATCTGAAGTATAGCTGCTAATCTGGAATTTGAAGCTCTCTAC 1077
Db 1162 CAAGATGGGATGAATCACTGACATACCAAGTTGCTGCTGAGAGCTGGGGCTCTTTAT 1221
Qy 1078 AC 1079
Db 1222 AC 1223

RESULT 12
CR607487 1870 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DK015YB21 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
CR607487
CR607487.1 GI:50488294
HTC; CDS/UT cDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1870)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segreff@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source Location/Qualifiers
1..1870

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC015YB21"
/issue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 14.5%; Score 167.6; DB 4; Length 1870;
Best Local Similarity 54.4%; Pred. No. 2.4e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

```
421 CCTAAGATTGTTGTCAGAGCATCGTGTCTATTATTGTGCGCTATAGAGTCGTGAA 480
574 CCGGAGGTTGTAGAGCCCGCTCCGAAAGCCATATTGTGCTCATGTGCGCGGAG 633
481 GCACATTGAGATTAATGCTCCCAATTGGCACTGCTGCTGCCAAACAACATTGAC 540
634 CACCACTGCGGCTGCTGCTTACCACTGACCCCTCTTGGAGCGCAGCAGCTGTGT 693
541 TATGCAATTTTCATTGTGAGCAAGTGGCAATCAGACGTTTAATGGCGGAACTAATG 600
694 TATGCACTTAATGTATCATCAACAGCTGGAATGAACTTTTACAGGGCAAACTGTG 753
601 AACGTTGATACGACCTAGCATCAACGCTCTTACCCATGCGAGCTTCACTTTTCATGAT 660
754 AACGTTGGGGTGGAGAGGCCCTGCGTGATGAAGATGGGACTGCTGTTCTTGACGAT 813
661 GTGCAATTACTGCGCGAAGATGACCTTAACCTGTACAGTGT---CGAATTCAACGAGT 717
814 GTGACCTCTTGGCAAAATGACCAATCTGTATGTGTGACCCCGGAGACCCCGC 873
718 CATATGAGTGAAGCATGATTAATCAATTATTAACCTTCCATATGCGGATCTTGGC 777
874 CATGTTGCCGTTGCTATGAACAAGTTGATACAGCTCCCGTACCCCAAGTACTTGGGA 933
778 GGAATCAGTCACTAACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTT 837
934 GGAAGTCTGACACTTAATCTCTGACAGTACCTGAAGATGAATGGCTTCCCAATGATAC 993
838 TGGGGTGGGGCGAGAGACGACGATTTGGCGACGAGAACTGATGCTGAGCTGA 897
994 TGGGGGCTGGGGTGGAGATGACGACATTTGCTACAGAGGTGCGCTGCGGAGTGAAG 1053
898 GTTTCAGATATCCGACAAATTTGACGATATTAATAATGATTAAGACTCGACGAGAGG 957
1054 ATCTCTCGGCCCCCATCTGTGAAGACATTAATAGTGTGAAGACGAGAGAGTAAAG 1113
958 ACGAATCCAGTTAATTAATCCGCTCAAAATTAATGGGCAAAAGAGCGCATGAGCA 1017
1114 GGCATAGAGAAATATCCACAGATTTGACCTCTGCTCCGTAACGAAATTCCTGAGG 1173
1018 CGTACGCGCTTAAGCAATTTGAATTAAGCTGTGAATCTGGAAATGGAAGCTCTTAC 1077
1174 CAAGATGGATGAATCACTGACATACGATGCTGCTGAGAGAGCTGGGGCTCTTTAT 1233
1078 AC 1079
1234 AC 1235
```

RESULT 13
CR592876 1876 bp mRNA linear HTC 21-JUL-2004
LOCUS CR592876
DEFINITION Full-length cDNA clone CSOD1008YCA21 of Placenta Cot 25-normalized
of Homo sapiens (human).

ACCESSION CR592876
VERSION CR592876.1 GI:50473683
KEYWORDS HTC; cDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1876)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue

REFERENCE 2 (bases 1 to 1876)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segreff@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source Location/Qualifiers
1..1876

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1008YCA21"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 14.5%; Score 167.6; DB 4; Length 1876;
Best Local Similarity 54.4%; Pred. No. 2.4e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

```
421 CCTAAGATTGTTGTCAGAGCATCGTGTCTATTATTGTGCGCTATAGAGTCGTGAA 480
575 CCGGAGGTTGTAGAGCCCGCTCCGAAAGCCATATTGTGCTCATGTGCGCGGAG 634
481 GCACATTGAGATTAATGCTCCCAATTGGCACTGCTGCTGCCAAACAACATTGAC 540
635 CACCACTGCGGCTGCTGCTTACCACTGACCCCTCTTACAGGCGCAGAGCTTGT 694
541 TATGCAATTTTCATTGTGAGCAAGTGGCAATCAGACGTTTAATGGCGGAACTAATG 600
695 TATGCACTTAATGTATCATCAACAGCTGGAATGAACTTTTACAGGGCAAACTGTG 754
601 AACGTTGATACGACGTGACATCAACCTCTTACCAATGGCAGTCTTATCTTTCAATGAT 660
755 AACGTTGGGGTGGAGAGGCCCTGCGTGATGAAGTGGGACTGCTGTTCTTGACGAT 814
661 GTGCAATTACTGCGCGAAGATGACCGTAACTGTACAGTGT---CGAATTCAACGAGT 717
815 GTGACCTCTTGGCAAAATGACCAATCTGTATGTGTGACCCCGGAGACCCCGC 874
718 CATATGAGTGAAGCATGATTAATTAATTAATTAATCTTCAATTTGGCGCATCTTGGC 777
875 CATGTTGCCGTTGCTATGAACAAGTTTGAATTAAGCTCTCCGTAACCCCAAGTACTTGGGA 934
778 GGAATCAGTCACTAACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTT 837
935 GGAAGTCTGACACTTAATCTGACAGTACCTGAAGATGAATGGCTTCCCAATGAATAC 994
```


ORIGIN /clone lib="riken1"
/note="CB inbred strain"

Query Match 14.3%; Score 165; DB 1; Length 673;
Best Local Similarity 56.6%; Pred. No. 1.2e-40;
Matches 306; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```
QY 393 GGACACGACGCGCGGTGACATGATGCTTAAGATTGTGTGCAAGCATCGTGTGC 452
    |||||
DB 542 GGACATTAAACGAGAGGACCTGGAACCCAAAGCTGTAGCCACGATGGAAGGTGC 483
    |||||
QY 453 TATTAATGACCCCTATGAGATCGTGAAGCACATTTGAGATATGCTCCACATTTGCA 512
    |||||
DB 482 GATCATCATTCCTTTTCGTATCGTACAGAGCATCTTCCATTTTCTTCGGCATCTGAT 423
    |||||
QY 513 CTGTTGCTCGCCAAACAACTGACTATGCAATTTTCATTTGAGACAAAGTGGCGAA 572
    |||||
DB 422 ACCGATGTTGCAGAAAGCAGCGGCTGGAATTTGCTCTATGTTGTTGAACAGACAGATAC 363
    |||||
QY 573 TCAGACGTTTATTCGCGGAAACTAATGAAGTTGATACGACGTACATCAAGCCTCTA 632
    |||||
DB 362 ACAACCTTTTATCGTACATGCTTTTAAACGTGCTTCAAGAGGCCATGAAGGATGT 303
    |||||
QY 633 CCCATGCGAGTGTCTCATCTTCATGATGTGATTTTACTGCCGAAGATGACCGTAACCT 692
    |||||
DB 302 TGCTGGGACCTGCAATATTTTCATGATGTGATCACTTACCTGAAATGACCGAAATTA 243
    |||||
QY 693 GTACAGTGTCCAAATTCACACGTCATATGAGTGTAGCGATCGATTAATTTCAATTATA 752
    |||||
DB 242 TTACGATGTGGAAGAAATGCGACGTCATTTTTCAGCAAAAGTTGACAAATACATGTACAT 183
    |||||
QY 753 ACTTCCATATTTGGGGGATTTGGGGGATCGATGCACTAACAAAAGATCACCTGAAGAA 812
    |||||
DB 182 TCTTCCATACAAATGAGTCTTGTGTGTGTAGTGAAGTGAACGATGGAACAAATTCAGAA 123
    |||||
QY 813 AATCAATGGAATTTTGAATGATTTTGGGGTGGGGCGGAGAGACGACGATTTGGCGAC 872
    |||||
DB 122 GATTATGGAATTTCCAAATGCTTCTGGGGTGGGGTGGAGAAAGATGATGATCTTTGGAA 63
    |||||
QY 873 GAGAACATGATGCTGCACTGAAGTTCGAAGATATCCGACACAATTTGACGATATATA 932
    |||||
DB 62 CAGGGTTCATCTATGCTGTGATACACGTAACAGACCAAGAGGAGACTTGGGAAATATACA 3
    |||||
QY 933 A 933
    |
DB 2 A 2
```

Search completed: January 8, 2006, 22:26:44
JDB time : 5197 secs

THIS PAGE BLANK (USPTO)